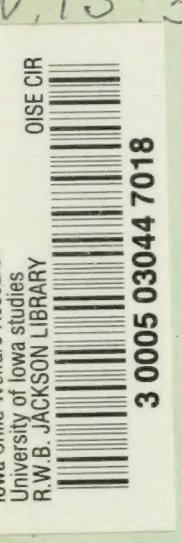
The Ontario Institute for Studies in Education Library



,no.3

THE LIBRARY

The Ontario Institute

for Studies in Education

Toronto, Canada

The Ontario Institute for Studies in Education Library

New Series No. 361 October 1, 1938

UNIVERSITY OF IOWA STUDIES

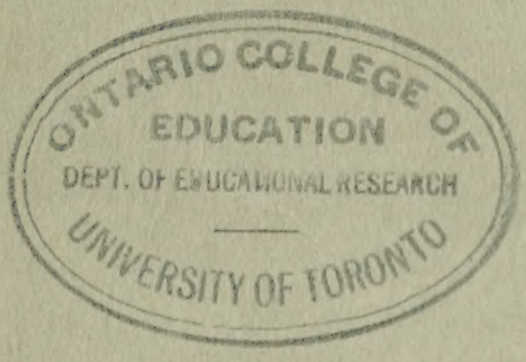
STUDIES IN CHILD WELFARE

Vol XV No. 3

THE INTELLIGENCE OF PRESCHOOL CHILDREN AS MEASURED BY THE MERRILL-PALMER SCALE OF PERFORMANCE TESTS

by

BETH L. WELLMAN, Ph.D.



Published by the University, Iowa City

Issued semi-monthly January 1 to June 30 and monthly July 1 to December 31. Entered at the post office at Iowa City, Iowa, as second class matter under the Act of October 3, 1917

THE INTELLIGENCE OF PRESCHOOL CHILDREN AS MEASURED BY THE MERRILL-PALMER SCALE OF PERFORMANCE TESTS

by

BETH L. WELLMAN, Ph.D.

GEORGE D. STODDARD, Ph.D., Editor
UNIVERSITY OF IOWA STUDIES
STUDIES IN CHILD WELFARE

LIBRARY
NOV 11 1970
THE ONTARIO INSTITUTE
FOR STUDIES IN EDUCATION

Volume XV, No. 3

ONTARIO COLLEGE OF EDUCATION
DEPT. OF EDUCATIONAL RESEARCH
UNIVERSITY OF TORONTO

Published by the University
Iowa City, Iowa
1938

ONTARIO COLLEGE OF EDUCATION
DEPT. OF EDUCATIONAL RESEARCH
UNIVERSITY OF TORONTO

FOREWORD

Dr. Wellman's monograph is the most intensive study of the Merrill-Palmer scale since Dr. Stutsman's original presentation in her book entitled *Mental Measurement of Preschool Children.* It may be read with profit by all persons who are interested not only in this test, but in the more general problem of mental measurement at the preschool levels.

On the basis of 510 testings of 281 children, it is clearly demonstrated that the Merrill-Palmer scale is not interchangeable with the Stanford-Binet. However it does provide a useful and interesting supplement to the latter, being less closely related to language facility and to the educational and occupational status of the parents. Its chief defect may lie in a dependence upon small time differences.

By careful analysis Dr. Wellman is able to differentiate between the practice effects commonly noted for the Merrill-Palmer scale and changes in scores indicative of increased brightness. Once more an effect of attendance in the preschools of the Station is shown to be a reliable acceleration in mental growth.

GEORGE D. STODDARD

Office of the Director
Iowa Child Welfare Research Station
University of Iowa
May 27, 1938

CONTENTS

Page

Setting for This Study ---- 7

Abilities Measured ---- 8

Problems Peculiar to Measurement of Preschool Children ---- 8

Technical Difficulties ---- 10

Subjects and Sources of Data ---- 11

Scoring Methods ---- 12
- Distribution of Scores ---- 13
- Distribution of Scores in Relation to Other Scoring Methods ---- 17
- Correlation Between Scorings ---- 19
- Relation to Chronological Age ---- 21
- Scores Used in This Study ---- 22

Change on Retests ---- 24
- Sigma Scores ---- 24
- Intelligence Quotients ---- 24
- Percentile Ranks ---- 26
- Changes on Retests for Constant Group ---- 26
- Retests at One Week ---- 29
- Explanation of Gains ---- 31
- Comparison With Scores on Minnesota Preschool Scale ---- 33
- Summary ---- 36

Effects of Preschool Attendance ---- 37
- Changes During Periods of Preschool Attendance and Non-attendance ---- 40
- Changes According to Number of Days in Attendance Between Tests ---- 57
- Changes With Length of Daily Session ---- 59
- Changes According to Interval Between Tests ---- 60
- Relation to Amount of Attendance Prior to First Test ---- 67
- Relation to Number of Days Preschool Attendance Prior to First Test ---- 73
- Summary ---- 75

Conclusions ---- 77

Socio-Economic Status ---- 78
- Occupation of Father ---- 78
- Education of Father ---- 85
- Education of Mother ---- 91
- Midparent Education ---- 92
- Summary ---- 93

Relation to Binet ---------- 95
Correspondence Between Merrill-Palmer and Binet on Tests One Week Apart ---------- 95
Changing Relationship After Preschool Attendance ---------- 100
Relation to Subsequent Binets ---------- 102
Summary ---------- 109

Comparison With Results of Other Investigations ---------- 111
Relation to Chronological Age ---------- 111
Sex Differences in Percentile Rank ---------- 112
Changes According to Interval Between Tests ---------- 113
Changes During Periods of Preschool Attendance and Non-attendance ---------- 122
Relation to Number of Days of Preschool Attendance Prior to Test ---------- 125
Correspondence Between Merrill-Palmer and Binet ---------- 127
Relation to Subsequent Binets ---------- 131
Relation to Minnesota Preschool Scale ---------- 131
Relation to Occupation of Father ---------- 133
Relation to Other Characteristics of Children ---------- 135
Prediction From Infant Tests ---------- 138
Comparison With English Norms ---------- 139
Summary ---------- 140

Summary of Results and Conclusions and an Evaluation of the Test ---------- 144

References ---------- 148

THE INTELLIGENCE OF PRESCHOOL CHILDREN AS MEASURED BY THE MERRILL-PALMER SCALE OF PERFORMANCE TESTS

The measurement of the intelligence of young children has become particularly challenging within recent years, due to the increased focus of attention upon these ages. The preschool child has become the object of an unprecedented number of investigations, in most of which it has been helpful to know the distribution of intelligence of the subjects. For practical purposes, such as for adoption and for recommendation of training programs, the desirability of intelligence measurements of young children has received increasing recognition. The necessity of a thorough trial of existing tests in order to learn their values, strengths, and weaknesses thus seems self-evident.

SETTING FOR THIS STUDY

That attendance at preschool laboratories increases the intelligence of young children when measured by the Kuhlmann and Stanford revisions of the Binet scale has been demonstrated by the author in two previously published articles (37, 38). More than six hundred preschool children were included in these studies. The gains made were shown to be real and not to be the result of practice effects (not to accrue from sheer repetitive experience with the test). The gains were not accounted for by occupational or educational status of the parents (5).

The higher level after gaining was maintained up through the age of fourteen years (the upper age limit of use of the Stanford-Binet with superior children) under certain school and group conditions. Under certain other school and group conditions, further significant gains were made. The effects of the gains were further shown in intelligence status at the senior high school levels as measured by the American Council of Education examination (40), and on college entrance examination. It appeared that there was a real and permanent advantage in producing gains during the preschool years.

The major aim of the present investigation was to study the effects of preschool attendance upon intelligence as measured by the Merrill-Palmer scale. Secondary aims arising out of the primary one were: to make a critical study of the Merrill-Palmer scale as an intelligence test for preschool children, with particular reference to suitability of

the scale for use in obtaining consecutive measurements, sex differences, and validity; to compare the Merrill-Palmer scale in the above respects with other scales, notably the Kuhlmann and Stanford revisions of the Binet scale, and to a limited extent, the Minnesota Preschool scale; and to compare our findings with those of other investigators.

ABILITIES MEASURED

Revisions of the Binet scale have frequently been designated as "verbal" types of tests, while the Merrill-Palmer scale is called a "performance" type of test. A cursory examination will reveal that the subtests of the two scales appear on the surface very different in content, although some subtests of the Binet scale can be classified as belonging to the performance type and some in the Merrill-Palmer scale as verbal. In spite of this general difference in appearance, it should be possible to use the two scales interchangeably, if they are equally valid measures of intelligence, and if each really has a high degree of validity.

PROBLEMS PECULIAR TO MEASUREMENT OF PRESCHOOL CHILDREN

There seems to be a rather prevalent feeling that intelligence tests for young children are less satisfactory than at the school ages. This feeling appears to have arisen from three main sources:

1. *The fact that the main intent in the standardization of the Stanford-Binet* (29) *was to provide a measuring instrument for children of school age.* The tests below five years in the scale were primarily intended to permit the adequate measurement of older children of below average intelligence.[1] Very few children younger than five years were used in the standardization process. This has often been used as an argument that the tests must therefore be inadequate, although this conclusion does not necessarily follow. The Kuhlmann-Binet (18), however, did make a directed attempt to provide suitable tests for much younger children, beginning its range at three months. The Merrill-Palmer scale (27) was designed especially for the preschool ages and none other.

2. *The fact that preschool age children who have been retested have characteristically shown considerable fluctuation in IQ.* The statement is often made that tests at the preschool ages are much less reliable than at the school ages. The so-called reliabilities that have

[1]This situation has been remedied in a newer standardization of the Stanford revision by Terman and Merrill (30). The results included in the study here reported are all based on the standardization published in 1916.

been obtained at the preschool ages have often covered a period of at least six months during which time the child's environment has been markedly altered by nursery school attendance. "Reliabilities" obtained under such conditions may reflect the effects of real changes in intelligence due to preschool experiences.

Goodenough's (8) correlations based on retests at six weeks of non-preschool children (children not attending nursery school) on the Kuhlmann-Binet ranged from .76 to .87 for the three age groupings, two, three, and four years, with a correlation of .82 for the total group of 300 children. These figures are not far below those ordinarily reported for school children (indeed, they overlap), but are not as high as the highest that has been reported for school age children. The correlation for Goodenough's total group is surprisingly close to the median correlation of .83 computed by Nemzek (24) from a survey of the extensive literature on constancy of the IQ of children of all ages on Stanford-Binets.

Stutsman's (28) correlation on retests on the Merrill-Palmer scale with an age range comparable to Goodenough's (3 to 5 years) and an interval fairly comparable (1 day to 1 month) is very similar to Goodenough's, the correlation between raw scores on retests for forty-six children being .85 ± .03. Raw scores are, however, related to chronological age and the correlation between retests is affected by the age range. When the age range was increased to 2 to 5 years with the interval two months or less, the correlation increased to .95 ± .01 for eighty-one children, and when the age range was narrowed to a 12 months' range, every child being between 3 and 4 years of age, and the interval was 1 to 11 months, the correlation dropped to .72 ± .05.

Our own data on Merrill-Palmer retests, reported more fully later in the study, show correlations on retests at one-week intervals ranging from .89 to .96 for different methods of scoring for forty-four children ranging in age from 28 to 60 months. The correlation between mental ages was highest, .96 ± .01, the other scoring methods eliminating the effect of chronological age and being respectively: sigma scores .92 ± .02, intelligence quotients .90 ± .02, and percentile ranks .89 ± .02.

3. *Confusion between accuracy of measurement at a given time and predictive value.* When dramatic environmental changes, such as attending a nursery school, result in increasing children's IQ's in direct ratio to the level of IQ on entrance to nursery school, the correlation between repeated tests is lowered. This does not necessarily mean that the child's status has not been adequately determined at each

testing. Disturbances in the relative placement of individuals within a group over a period of time need not be attributed entirely to errors of measurement. They may represent real change.

Technical Difficulties

There are certain technical difficulties in the construction of intelligence tests for young children that work in the direction of reducing reliability. Each subtest carries a greater weight of the total IQ at the younger ages than at the older ages. Due to the smaller chronological age as denominator, the passing or failing of one test item will change the IQ several points. An illustration will make this clear. A child of 42 months who has a mental age of 42 months receives an IQ of 100. If the child had passed one more item on the Stanford revision, his mental age would have been 44 months and his IQ 105. A child of 48 months, by passing an additional item would have increased his IQ by 3 points, and one of 84 months by 2 points.

On the Kuhlmann revision, which has eight subtests in each year, an additional item passed adds 1.5 months of mental age. At 24 months, one extra item would increase the IQ from 100 to 106, at 30 months to 105, and at 36 months to 104.

On the Merrill-Palmer scale, each item passed is not given mental age credits as in the Binet scales, but the total point score can be credited with a mental age equivalent. Mental age is one of several different methods of scoring, all of which are based upon and derived from the total point score.

Inconsistencies in IQ can arise at these young ages out of inconsistencies in retaining or rejecting decimals of months. A child of 2 years, 0 months and 14 days may be considered as 24.0 months or 24.5 months of age. If his mental age is 25.5, the former age will yield an IQ of 106 and the latter 104. If the decimal point is dropped in mental age also, the IQ will then be 104, providing the mental age is considered as 25.0 months; or 108, providing the mental age is considered as 26.0 months.

In spite of these difficulties, the Binet IQ's of groups of preschool children do change in a significant and orderly manner with attendance at preschool laboratories. Do the children also show systematic orderly changes when measured by the performance type of test, more specifically by the Merrill-Palmer scale? Is the Merrill-Palmer as adequate a measure of intelligence as the Binet tests? Can it be used interchangeably with them? Does a combination of "verbal" and "performance" tests yield a better measure than either alone?

SUBJECTS AND SOURCES OF DATA

Five hundred ten Merrill-Palmer tests were administered to 281 children 20 to 62 months of age. Two hundred eight children were in attendance at the preschool laboratories of the Iowa Child Welfare Research Station at Iowa City, Iowa; eleven were in attendance at the nursery school of the Iowa State Teachers' College at Cedar Falls, Iowa; seventeen were residents of the City of Childhood at Mooseheart, Illinois; and forty-five were residents at the Iowa Soldiers' Orphans' Home at Davenport, Iowa.

The tests and retests were distributed as follows:

	Test						
	First	Second	Third	Fourth	Fifth	Sixth	Total
Preschool Laboratories	208	90	47	22	9	2	378
Cedar Falls	11	6	2				19
Mooseheart	17						17
Orphan's Home	45	31	19	1			96
Total	281	127	68	23	9	2	510

All tests were given by experienced examiners who were, with one exception, research assistants in the Iowa Child Welfare Research Station; this examiner was a graduate laboratory assistant who supervised the giving of tests in connection with a course in clinical psychology. The children were examined in familiar surroundings. A special effort was made to see that the children became acquainted with the examiners prior to the time of examination. Preschool laboratory children were allowed a preliminary period in which to become adjusted to the preschool situation; at least two weeks, and usually more, were allowed after entrance to preschool before any child was tested. Some of the children had their first test after several months of preschool attendance. An effort was made to test each child not more than a week away from the day of the month on which his birthday fell. This was done in order to avoid fractions of months in chronological age.

The distribution of chronological ages at the time of the first test was as follows:

Age, Months	Total Group	Preschool Laboratories Only
20 to 22	2	0
23 to 25	10	10
26 to 28	17	15
29 to 31	19	14
32 to 34	30	24
35 to 37	34	29
38 to 40	21	15

Age, Months	Total, Group	Preschool Laboratories Only
41 to 43	30	17
44 to 46	26	18
47 to 49	23	18
50 to 52	25	21
53 to 55	18	13
56 to 58	20	14
59 to 61	5	2
62 to 64	1	0
Total	281	208

Figure 1 also presents the distribution of ages. The mean chronological age for the total group was 41.4 months, S.D. 9.7 and for the preschool laboratory group 40.8 months, S.D. 9.7.

The group as a whole was somewhat superior mentally, as is shown by the mean scores given in Table 6. Mental age of the total group on the first test was 46.4, 5.0 months in advance of chronological age, and of the preschool laboratory group 48.2, 7.4 months in advance of their chronological age.

SCORING METHODS

Stutsman (27) provides for the following methods of scoring the results of the test: raw score, mental age equivalent of score, standard deviation in terms of score value (called hereafter sigma score), percentile rank in terms of score value and intelligence quotient. In

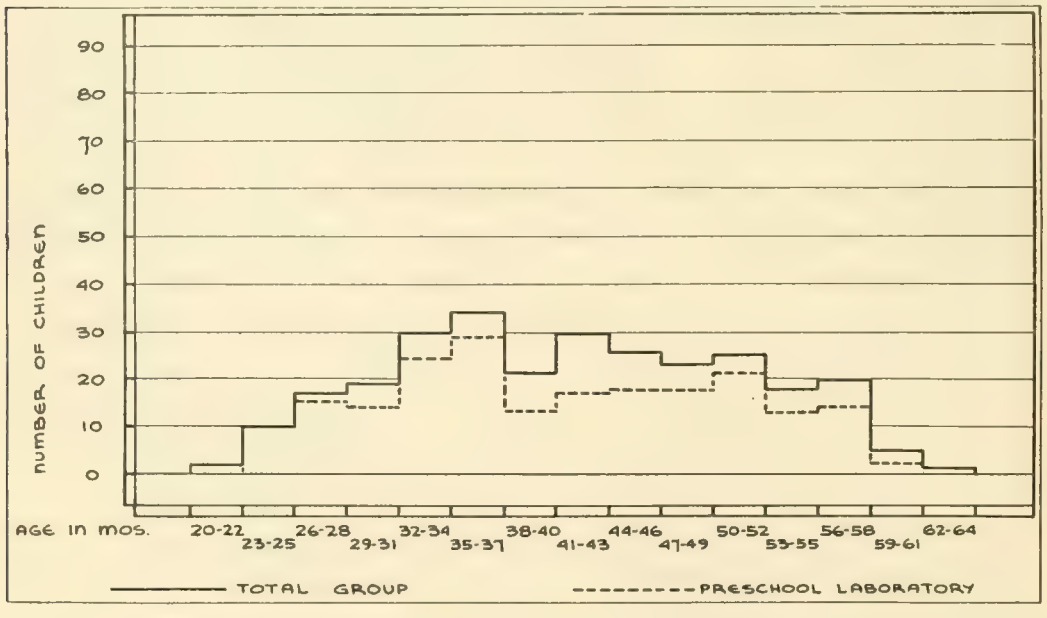


Figure 1. Distribution of Chronological Ages of Total Group and of Preschool Laboratory Group at Time of First Test

addition, she gives tables of standard deviations in terms of mental age for chronological age groups by months, and standard deviations in terms of intelligence quotient for chronological age groups.

Stutsman states: "For purposes of comparing the performance of children of different ages, or the performance of the same child at different ages, interpretation of the score by standard deviation in terms of score value or mental-age value and interpretation by percentile rank are preferable to interpretation in terms of the intelligence quotient." (p. 235) She points out that the same IQ at different ages places the child very differently in a distribution. "An IQ of 119 may have a range of values at different age levels from $+0.5$ to $+2.0\sigma$. In other words, a child of 27 months of age who has an IQ of 119 has attained the same level of superiority as a child of 48 months of age who has an IQ of 154. Interpretation of the score in terms of the intelligence quotient assumes that the variability in terms of mental age increases at an evenly progressive rate with zero as an origin. Although this should possibly be true with an ideal scale, it is not true of the present Merrill-Palmer scale." (p. 235)

Distribution of Scores

In order to decide which of the various methods of scorings to use in this study, a distribution was made of the scores by each method on the 281 first tests. These are given in Figure 2; comparable distributions for the 208 preschool laboratory children are also given in Figure 2.

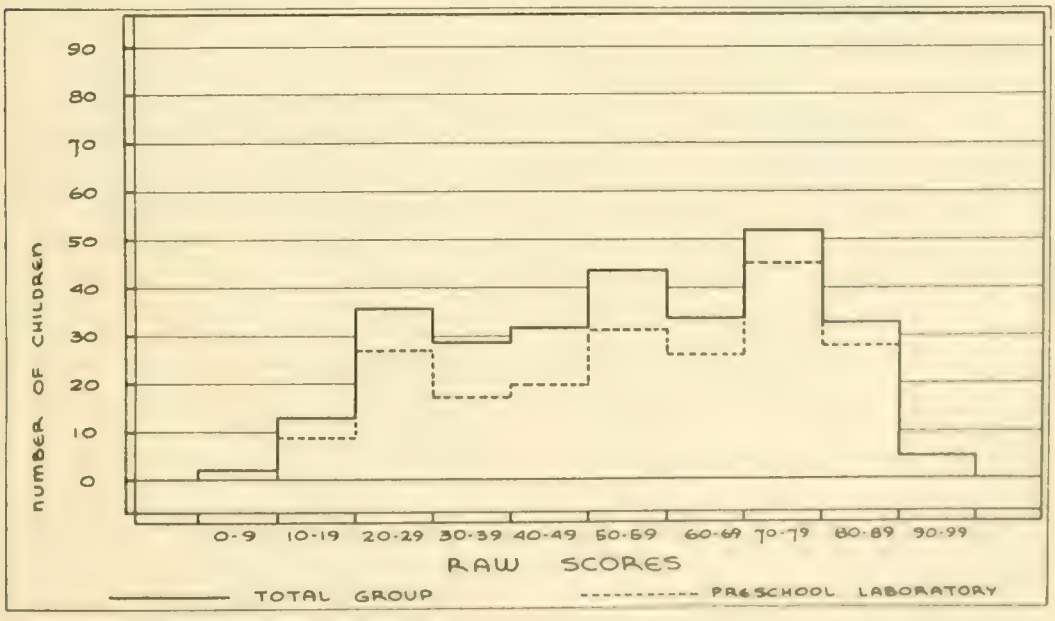
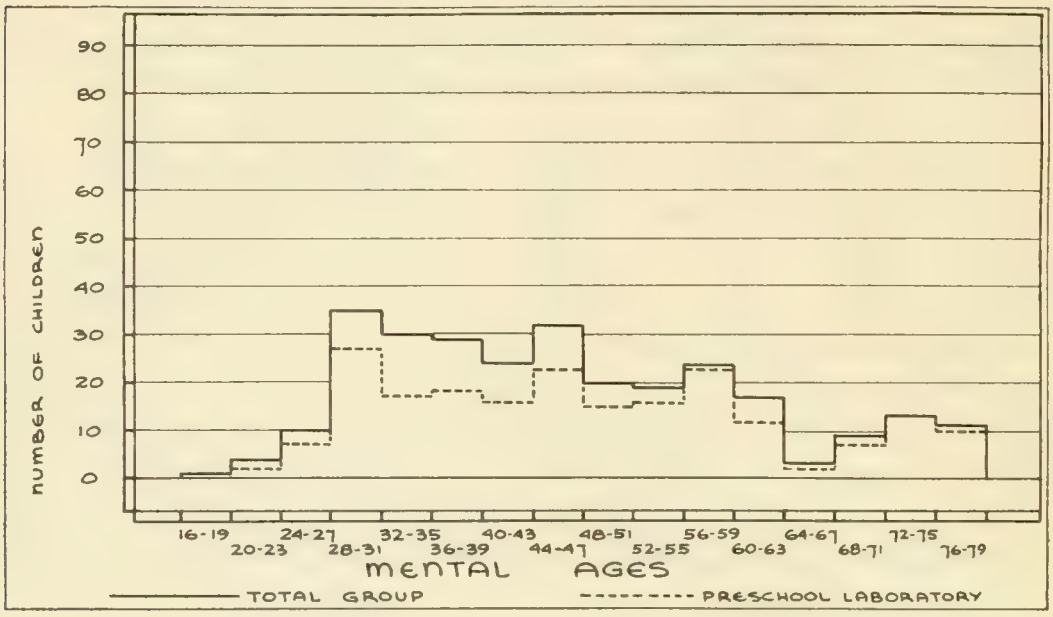
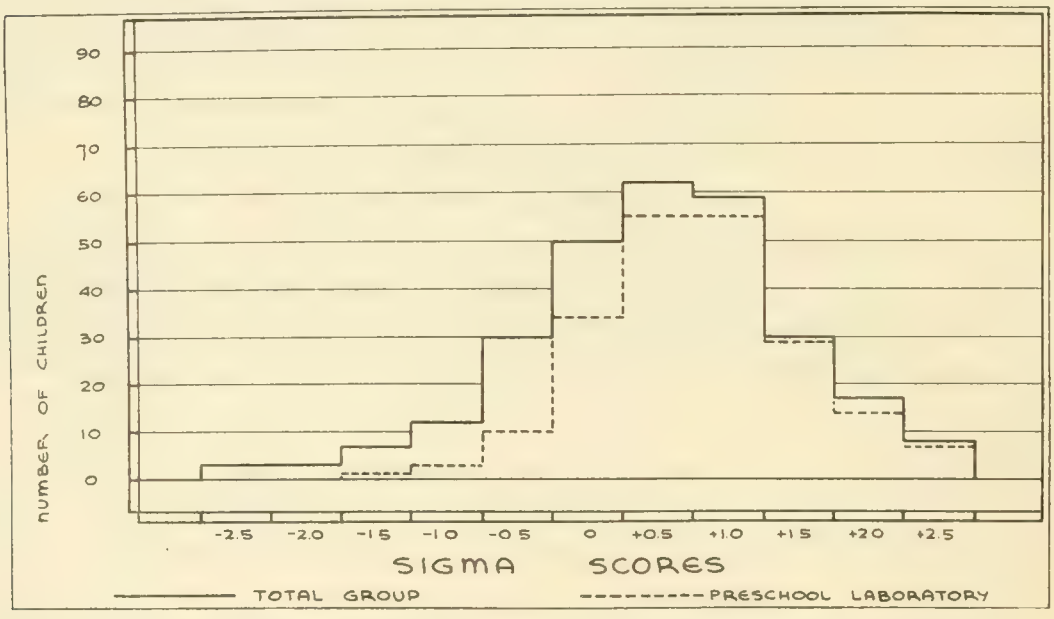


Figure 2. Distribution of Scores on First Test.





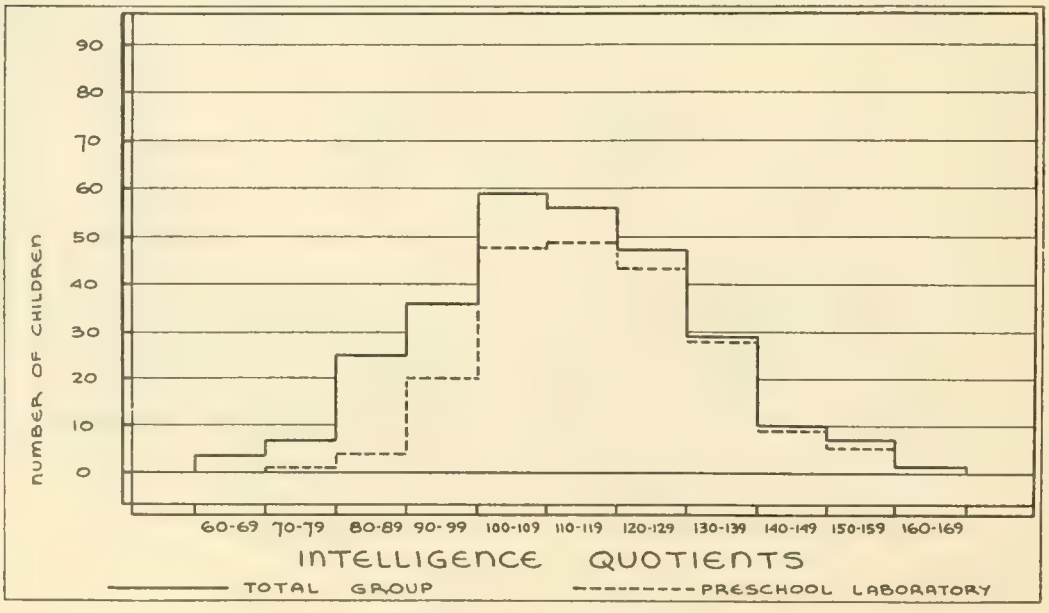


Figure 2. (Continued)

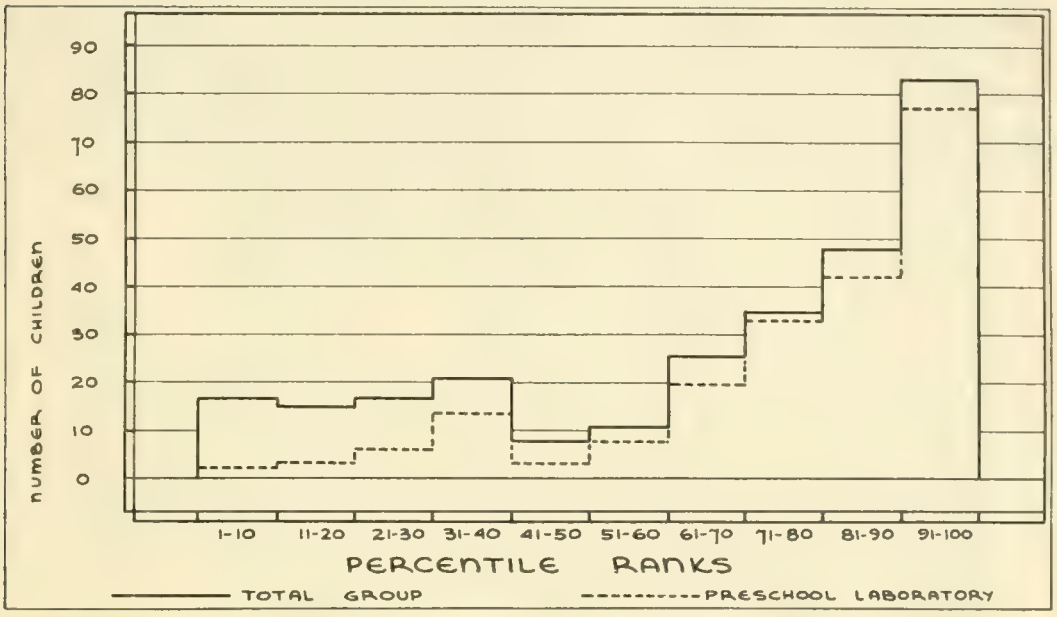


Figure 2. (Continued) Distribution of Scores on First Test.

The raw scores on first test in our data were distributed as follows:

Score	Total Group	Preschool Laboratories
0 to 9	2	0
10 to 19	13	9
20 to 29	36	27
30 to 39	29	17
40 to 49	32	20
50 to 59	44	31
60 to 69	34	26
70 to 79	52	45
80 to 89	33	28
90 to 99	5	5

The distribution of mental ages was as follows:

Mental Age, Months	Total Group	Preschool Laboratories
16 to 19	1	0
20 to 23	4	2
24 to 27	10	7
28 to 31	35	27
32 to 35	30	17
36 to 39	29	18
40 to 43	24	16
44 to 47	32	23
48 to 51	20	15
52 to 55	19	16
56 to 59	24	23
60 to 63	17	12
64 to 67	3	2
68 to 71	9	7
72 to 75	13	13
76 to 79	11	10

Sigma scores in terms of raw scores were read from Table 27, page 237, according to the instructions given by Stutsman (27). No interpolations were made. If a score fell between two sigma values of the table, it was assigned the sigma value which was closer to zero.

The sigma scores were distributed as follows:

Sigma Score	Total Group	Preschool Laboratories
—2.5	3	0
—2.0	3	0
—1.5	7	1
—1.0	12	3
—0.5	30	10
0	50	34
0.5	62	55
1.0	59	55
1.5	30	29
2.0	17	14
2.5	8	7

The intelligence quotients were distributed as follows:

Intelligence Quotient	Total Group	Preschool Laboratories
60 to 69	4	0
70 to 79	7	1
80 to 89	25	4
90 to 99	36	20
100 to 109	59	48
110 to 119	56	49
120 to 129	47	43
130 to 139	29	28
140 to 149	10	9
150 to 159	7	5
160 to 169	1	1

Percentile ranks in terms of score values per age in months were read from Table 30, page 240, in Stutsman (27), interpolation being made to the nearest whole percentile. When scores fell beyond the 99th percentile of the table, a rank of 100 was assigned. The range of ranks assigned in this group was 1 to 100. The distribution of percentile ranks was:

Percentile Rank	Total Group	Preschool Laboratories
1 to 10	17	2
11 to 20	15	3
21 to 30	17	6
31 to 40	21	14
41 to 50	8	3
51 to 60	11	8
61 to 70	26	20
71 to 80	35	33
81 to 90	48	42
91 to 100	83	77

The intelligence quotients ranged from 61 to 162, with a mean of 111.4 for the total group and 116.7 for the preschool laboratory group. There was a fairly even distribution on either side of the mean. Percentile ranks varied from 1 to 100, with a mean of 68.0 for the total group and 78.0 for the preschool laboratory group, in a skewed distribution. More than one-third (37 per cent) of the preschool laboratory children received percentile ranks above 90, and one-fourth (26 per cent) above 95 on their first test. In a study of repeated measures on the same individuals within such a group as the preschool laboratory, the method of scoring in terms of the percentile ranks offered by Stutsman presents difficulties because the ranks do not differentiate among high performances and hence adequately allow for registering of increases. At the other end of the distribution the same problem is met with regard to possible decreases. In the total group a number of children received percentile ranks of 1 to 10. Of the children receiving percentile ranks of 1 to 5, four had intelligence quotients falling in the step interval 60 to 69, five in the step interval 70 to 79, and two in the interval 80 to 89. Two children receiving percentile ranks of 6 to 10 had intelligence quotients 70 to 79 and four 80 to 89. That the intelligence quotients were not unduly high for these subjects is confirmed by Kuhlmann-Binets obtained within one week of the Merrill-Palmers on four of the children whose percentile ranks were 1, 5, 5, and 5 respectively. Their Merrill-Palmer IQ's were 61, 78, 80, and 85, while their Kuhlmann-Binet IQ's were 59, 96, 77, and 97 respectively.

Distribution of Scores in Relation to Other Scoring Methods

The distribution of intelligence quotients per percentile rank for the total group is given in Table 1. It will be seen that the correspondence between the two methods of scoring is quite close in the middle ranges of percentiles. By way of contrast, the distribution of Kuhlmann- and Stanford-Binets per percentile rank is given in Table 2 for all children who had both Merrill-Palmer and Binet tests within one week.

While, with the exception of the 91 to 100 percentile group, the Merrill-Palmer intelligence quotients for a ten-percentile step interval are concentrated in one, two, or three intelligence quotient groupings, the Binet quotients are spread in some instances to cover nine group-

Table 1

Distribution of Merrill-Palmer Intelligence Quotients by Percentile Rank

Percentile Rank	Intelligence Quotient										
	60 to 69	70 to 79	80 to 89	90 to 99	100 to 109	110 to 119	120 to 129	130 to 139	140 to 149	150 to 159	160 to 169
1 to 10	4	7	6								
11 to 20			15								
21 to 30			4	13							
31 to 40				20	1						
41 to 50				3	5						
51 to 60					11						
61 to 70					23	3					
71 to 80					19	14	2				
81 to 90						34	14				
91 to 100						5	31	29	10	7	1
Total	4	7	25	36	59	56	47	29	10	7	1

ings, or ninety points in IQ. The eleven children receiving percentile scores of 51 to 60 all received intelligence quotients of 100 to 109 on the Merrill-Palmer scale. The seven children receiving percentile scores of 51 to 60, to whom Binet tests were given, had Binet intelligence quotients distributed from the step interval 80 to 89 to the step interval 130 to 139. Twenty-three of the twenty-six children receiving percentile ranks of 61 to 70 received Merrill-Palmer intelligence quotients of 100 to 109, while the twenty children receiving these percentile ranks who were given the Binet test were distributed in Binet intelligence quotients from the step interval 100 to 109 to the step interval 160 to 169. It is to be expected that the Merrill-Palmer intelligence quotients will agree more highly with the percen-

Table 2

Distribution of Kuhlmann and Stanford-Binet Intelligence Quotients by Percentile Rank on Merrill-Palmer

Percentile Rank	Intelligence Quotient												
	50 to 59	60 to 69	70 to 79	80 to 89	90 to 99	100 to 109	110 to 119	120 to 129	130 to 139	140 to 149	150 to 159	160 to 169	170 to 179
1 to 10	1	0	1	0	2								
11 to 20			2	1	1	1	0	1					
21 to 30				1	3	2	2	1					
31 to 40					1	7	2	2					
41 to 50					1	1	1	1	0	1			
51 to 60				1	0	0	1	2	3				
61 to 70						5	6	5	3	0	0	1	
71 to 80					3	7	4	8	2	2			
81 to 90			1	0	0	3	6	12	10	6	2		
91 to 100					3	8	8	16	14	10	9	0	2
Total	1	0	4	3	14	34	30	48	32	19	11	1	2

tile ranks than the Binet results, since both percentiles and intelligence quotients are derived from the same raw scores. The relationship to Binet results will be discussed more fully later in this study.

The children who received percentile ranks of 91 to 100 were distributed in intelligence quotient from the step interval 110 to 119 to the step interval 160 to 169.

Sigma score means were +0.49 for the total group and +0.78 for the preschool laboratory group. There was a fairly even distribution of scores on either side of the mean, there being no concentration of scores at the upper end of the distribution as in percentile ranks. The distribution of sigma scores per percentile rank is given in Table 3. It

Table 3

Distribution of Sigma Scores by Percentile Rank

Percentile Rank	Sigma Score										
	—2.5	—2.0	—1.5	—1.0	—0.5	0.0	0.5	1.0	1.5	2.0	2.5
1 to 10	3	3	7	4							
11 to 20				8	7						
21 to 30					16	1					
31 to 40					7	14					
41 to 50						8					
51 to 60						11					
61 to 70						16	9	1			
71 to 80							34	1			
81 to 90							19	27	2		
91 to 100								30	28	17	8
Total	3	3	7	12	30	50	62	59	30	17	8

will be seen that children who received the lowest percentile ranks (1 to 10) were spread over a considerable portion of the sigma score range, and that those receiving percentile scores above 90 received sigma scores varying from +1.0 to +2.5.

The distribution of sigma scores per intelligence quotient is given in Table 4.

Correlation Between Scorings

The degree of correspondence between the different methods of scoring is expressed by correlations in Table 5. There is no greater correspondence between the two methods recommended by Stutsman (percentile and sigma score) than between intelligence quotient and either of these. The most uniform array of correlations is between intelligence quotient and sigma score. The correlation tends to be

Table 4
Distribution of Sigma Scores by Intelligence Quotient

Intelligence Quotient	Sigma Score										
	—2.5	—2.0	—1.5	—1.0	—0.5	0.0	0.5	1.0	1.5	2.0	2.5
60 to 69	3	1									
70 to 79		2	4	1							
80 to 89			3	11	11						
90 to 99					19	17					
100 to 109						33	26				
110 to 119							28	26	2		
120 to 129							8	20	9	8	2
130 to 139								12	10	6	1
140 to 149								1	7	0	2
150 to 159									2	3	2
160 to 169											1

Table 5
Correlations by Age Groups Between Scorings on First Tests

Age, Months	Children	Percentile and Sigma Score		Percentile and Intelligence Quotient		Sigma Score and Intelligence Quotient	
		r	Probable Error	r	Probable Error	r	Probable Error
20 to 29	34	.89	.02	.94	.01	.96	.01
30 to 41	110	.89	.01	.89	.01	.96	.01
42 to 53	97	.95	.01	.92	.01	.95	.01
54 to 62	40	.93	.01	.98	.01	.96	.01
20 to 62	281	.90	.01	.90	.01	.90	.01

lower when the total age range is included than when the age range is narrowed. The explanation appears to lie in some peculiarities in the derivation of these scores from raw score in the original standardization group. In which method of scoring the difficulty lies is not apparent from the data. While the range of sigma scores is relatively small for a given IQ level within an age group, it becomes larger for the total group because the actual sigma scores included differ from age to age. The following tabulation of sigma scores for the IQ level 120 to 129 will illustrate this point.

Age, Months	Sigma Scores				
	0.5	1.0	1.5	2.0	2.5
20 to 29			1	5	2
30 to 41		4	8	3	
42 to 53		14			
54 to 62	8	2			
20 to 62	8	20	9	8	2

It seems there is some disjunction between mental age derivations and sigma score derivations to produce this type of distribution. The

above type of distribution is not peculiar to IQ in relation to the other scores, however, as is implied by Stutsman. It holds for percentile ranks in relation to sigma scores as well, as will be seen in the following distribution of sigma scores for children who received percentile ranks of 91 to 100.

Age, Months	Sigma Scores			
	1.0	1.5	2.0	2.5
20 to 29		1	5	2
30 to 41	7	9	9	6
42 to 53	13	17	3	
54 to 62	10	1		
20 to 62	30	28	17	8

The distribution of intelligence quotients for the children receiving percentile scores of 91 to 100 was as follows:

Age, Months	Intelligence Quotients					
	110 to 119	120 to 129	130 to 139	140 to 149	150 to 159	160 to 169
20 to 29		8				
30 to 41	5	13	8	2	2	1
42 to 53		10	12	6	5	
54 to 62			9	2		
20 to 62	5	31	29	10	7	1

Relation to Chronological Age

Chronological age correlated with each of the scoring methods to the following extent:

Scoring Method	Children	*r*	P.E.
Percentile	281	.26	± .04
Sigma Score	281	—.11	± .04
Intelligence Quotient	281	.12	± .04
Raw Score	281	.82	± .01

There was some tendency for the older children to receive the higher percentile ranks, indicated by the correlation of .26 ± .04. Sigma scores and intelligence quotients showed little relationship to age.

Mental age was correlated with chronological age as follows:

Chronological Age, Months	Children	*r*	P.E.
20 to 29	34	.68	± .06
30 to 41	110	.45	± .06
42 to 53	97	.51	± .05
54 to 62	40	— .07	± .01
20 to 62	281	.81	± .01

The size of the correlation between mental age and chronological age varied considerably between age groups. There appears no ready explanation for the lack of relationship at the five-year level, other

than the lack of ceiling of the test. The correlation may have been lowered somewhat by the fact that the oldest children in the group were orphanage children of relatively lower ability. No preschool laboratory child over fifty-nine months of age was included. The range of ages within the preschool laboratory group was too small to permit a separate correlation at this age level. Of the twenty-six preschool cases, nineteen received mental ages of 70 to 79 months. The number of cases (14) of other children at this age level was too small for adequate separate correlation.

The correlation between mental age and chronological age for the total group, .81 ± .01, is considerably below the correlation of .90 reported by Bristol (4) for 531 English children and the correlation of .92 ± .004 reported by Stutsman (27, p. 116) between chronological age and raw score for 631 cases. It corresponds more closely with Goodenough's (9) correlation of .85 between mental age and chronological age for thirty-three nursery school children, ages 27 to 59 months of age. Raw score in our data correlated with chronological age .82 ± .01. Since our group was probably more highly selected than Stutsman's or Bristol's, a somewhat lower correlation might be expected. Our correlation is very close to the one reported by Kawin (15) for 169 children 23 to 57 months of age drawn from various sources. Her correlation was .79.

Scores Used in This Study

In spite of Stutsman's warning against the use of the intelligence quotient, it has certain advantages. Stutsman says (27, p. 116): "A careful inspection of these tables will show that it is impracticable to use the intelligence quotient with the Merrill-Palmer test scale. At the different chronological-age levels the range of IQ's at —2.5 σ varies from 58 to 70. At —2.0 σ it varies from 66 to 79, and at —1.5 σ from 74 to 83. The nearer 0 σ the score falls, the less is the variation between the different age levels. At the other end of the distributions the variation is even greater. At +2.5 σ the IQ varies from 122 to 165; at +2.0 σ, from 119 to 154; and at +1.5 σ, from 114 to 141. It is apparent that with this amount of variation such an index has little significance. Accordingly, an interpretation in terms of the standard-deviation value of the score attained or in terms of percentile rank has a greater validity in the Merrill-Palmer scale than has the intelligence quotient."

Both the sigma score and intelligence quotient are dependent upon the raw score. The intelligence quotient depends directly upon the mental age equivalents assigned to the raw scores.[2] Stutsman's argument might have turned against the sigma score as justifiably as against the IQ in that the sigma score failed to differentiate at the extremes of the distribution where the IQ did differentiate. Comparisons between sigma scores at different age levels also assume that the distributions of mental levels (brightness) were comparable at the different ages in the standardization.

The intelligence quotient possesses the advantage of direct comparison with the results of other tests and of differentiation between children at all points in our distribution. For this reason, and others previously discussed, it was decided to retain this method of scoring throughout this study. Although it seemed probable from inspection of the distribution of percentile ranks on first tests that because of the large number of cases falling in the extreme upper end of the distribution increases might not be registered, the results in Table 6 do indicate changes. Hence it was decided to carry through this method of scoring in other analyses. Throughout this study, therefore, analyses are made on the basis of the three methods of scoring, sigma score, percentile rank, and intelligence quotient.

Table 6

Mean Chronological Ages, Sigma Scores, Intelligence Quotients and Percentile Ranks on Merrill-Palmer Retests, With Their Standard Deviations

Test	Children	Chronological Age		Sigma Score		Intelligence Quotient		Percentile Rank	
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Total Group									
First	281	41.4	9.7	0.49	0.96	111.4	19.0	68.0	30.0
Second	127	43.3	8.7	0.83	0.97	118.7	21.4	75.1	28.0
Third	68	48.2	7.8	0.65	0.81	117.0	18.0	75.2	26.7
Preschool Laboratory									
First	208	40.8	9.7	0.78	0.74	116.7	15.7	78.0	22.2
Second	90	42.1	8.2	1.10	0.81	123.6	19.1	82.9	22.0
Third	47	46.0	7.3	1.00	0.59	124.3	13.6	87.3	14.3
Fourth	22	47.4	5.7	1.38	0.50	136.6	11.3	95.3	5.1
Fifth	9	52.4		1.28		136.7		95.3	
Sixth	2	55.5		1.50		140.5		98.0	

[2]It is interesting to note the high degree of correspondence between our results and those of Driscoll (7) who derived the intelligence quotients used in her study by converting the mental ages into standard deviation values which, in turn, were converted into appropriate intelligence quotient values. Driscoll states that the difference between the IQ obtained in this way and one obtained directly by dividing the mental age by the chronological age ranged from 1 to 5 points. Comparisons with Driscoll's results are discussed more fully in the section of this study pertaining to comparisons with the results of other investigators.

CHANGE ON RETESTS

One hundred twenty-seven children were given a second Merrill-Palmer test and sixty-eight children were given a third test. The mean sigma scores, intelligence quotients, and percentile ranks on each test, with their standard deviations, are given in Table 6. The results for the preschool laboratory portion of the group are given separately in the table. Ninety preschool laboratory children received a second test, forty-seven a third, twenty-two a fourth, nine a fifth, and two a sixth test. Reasons for retest varied in accordance with the purpose of the particular section of the study under investigation at the time. All readily available children meeting the requirements for a given section were tested. Retests on preschool laboratory children were largely determined by continued enrollment in the laboratories. It will be seen from Table 6 that the retest scores were higher throughout than first test scores.

Sigma Scores

The mean sigma score on first test was 0.49 with a standard deviation of 0.96. On second test the mean increased to 0.83, but on third test dropped to 0.65. The preschool laboratory group was higher than the total group on the first test, having a mean of 0.78, which rose to 1.10 on second test but did not increase on third test. On fourth test the mean increased again, reaching 1.38. Because the retest groups represented only a portion of the original first test group, the statistical significance of the differences was not computed. In Table 7, which gives the means for the identical groups of individuals on each test, the ratios of the differences to the standard deviations of the differences are presented.

Intelligence Quotients

The mean intelligence quotient for the entire group on first test was 111.4 (Table 6). On second test it rose to 118.7, but on third test it remained about the same as on second test, the mean being 117.0. In the preschool laboratory group there was a difference of 6.9 points between first and second tests, no difference between second and third

Table 7

Mean Chronological Ages, Sigma Scores, Intelligence Quotients and Percentile Ranks, with Their Standard Deviations of Preschool Laboratory Children Retested on Merrill-Palmer

Test	Children	Chronological Age, Months	Sigma Score		Intelligence Quotient		Percentile Rank	
			Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Two Test Group								
First	90	36.7	.96	.82	113.3	15.1	74.2	24.9
Second	90	42.1	1.10	.81	123.6	19.1	82.9	22.0
Difference		5.4	.14		10.3		8.7	
Ratio of difference to standard deviation of difference			1.17		4.0		2.5	
Three Test Group								
First	47	34.3	.76	.88	112.1	15.2	71.5	25.9
Second	47	39.9	1.11	.81	121.5	17.1	82.6	20.9
Third	47	46.0	1.00	.59	124.3	13.6	87.3	14.3
Difference								
First and second		5.6	.35		9.4		11.1	
Second and third		6.1	—.11		2.8		4.7	
Ratio of difference to standard deviation of difference								
First to second			2.06		2.8		2.3	
Second to third			.73		.9		1.3	
Four Test Group								
First	22	29.9	1.02	.83	115.0	13.9	78.8	21.7
Second	22	35.5	1.41	.82	123.6	15.7	86.9	19.2
Third	22	41.9	1.20	.51	127.1	13.1	91.7	7.8
Fourth	22	47.4	1.38	.50	136.6	11.3	95.3	5.1
Difference								
First and second		5.6	.39		8.6		8.1	
Second and third		6.4	—.21		3.5		4.8	
Third and fourth		5.5	.18		9.5		3.6	
Ratio of difference to standard deviation of difference								
First to second			1.56		1.9		1.3	
Second to third			1.00		.8		1.1	
Third to fourth			1.20		2.6		1.8	

tests, but another large difference (12.3 points) between third and fourth tests. This meant that there was approximately 20 points superiority of the fourth test group over the first test group.

Percentile Ranks

Percentile ranks [3] showed practically the same results as intelligence quotients, with a gain from 68.0 on first test to 75.1 on second test for the total group, and no change from second to third test. In the preschool laboratory group there was a steady increase from 78.0 on first test to 95.3 on fourth test, a difference of 17.3 percentile ranks from first to fourth test.

Changes on Retests for Constant Group

Since not all children were retested, there was a smaller number of cases with each succeeding test. The means in Table 6 might, therefore, be influenced by the selection of cases for retests. Consequently, means on successive tests for the same individuals were computed. Table 7 gives the means on first and second tests for preschool laboratory children having two tests, the means on first, second, and third tests for preschool laboratory children having three tests, and the means on first, second, third, and fourth tests for the children having four tests. Table 8 presents similar material for orphanage children.

Sigma Scores

The mean sigma scores of the ninety preschool laboratory children having two tests was 0.96 on first test and 1.10 on second test, an increase of 0.14 over an interval of 5.4 months. The increase was not statistically significant, the ratio of the difference to the standard deviation of the difference being 1.17, or 87 chances in 100 of a true difference.[4]

The forty-seven preschool laboratory children having three tests

---

[3]In this study percentile ranks, whether on first test or retest, were all derived from Table 30, page 240 in Stutsman (27). This procedure was followed in order to register whatever changes might occur. In the volume *The Merrill-Palmer Standards of Physical and Mental Growth* (41), separate percentile tables are given for first tests and retests, since it had been noted that retest results were higher. According to these tables, at the chronological age of 40 months, the fiftieth percentile was a score of 51 on first test and 60 on retests. At the chronological age of 45 months, the fiftieth percentile was a score of 60 on first test and 69 on retests.

[4]Garrett, Henry E.: Statistics in Psychology and Education. New York: Longmans, Green, [c. 1926]. Pp. xiii, 317. (p. 134)

Table 8

Mean Chronological Ages, Sigma Scores, Intelligence Quotients and Percentile Ranks, with Their Standard Deviations, of Orphanage Children Retested on Merrill-Palmer

Test	Children	Chronological Age, Months	Sigma Score		Intelligence Quotient		Percentile Rank	
			Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Two Test Group								
First	30	46.0	—.55	.96	92.4	18.3	32.1	22.7
Second	30	46.3	—.02	.91	102.8	20.7	50.6	30.6
Difference		.3	.53		10.4		18.5	
Ratio of difference to standard deviation of difference			2.21		2.1		2.7	
Three Test Group								
First	18	46.6	—.58	1.04	92.4	19.8	30.6	28.0
Second	18	46.8	—.02	.96	102.1	24.2	47.4	33.0
Third	18	53.3	—.25	.65	97.6	16.2	42.8	25.0
Difference								
First and second		.2	.56		9.7		16.8	
Second and third		6.5	—.23		—4.5		—4.6	
Ratio of difference to standard deviation of difference								
First to second			1.70		1.3		1.7	
Second to third			.85		.7		.5	

showed a mean sigma score on first test of 0.76, on second test 1.11, and on third test 1.00. The intervals between tests were respectively 5.6 months and 6.1 months. The gain of 0.35 between first and second test yielded a ratio of 2.06, or 98 chances in 100 of a true gain greater than zero. From second to third test there was a slight loss in the mean, but it was not a significant loss, the ratio being .73 and the chances in 100 that it was a true loss being 76.

The twenty-two children who had four tests gained 0.39 sigma score from first to second tests, lost 0.21 from second to third, and gained 0.18 from third to fourth tests. None of these changes was significant, the ratios being respectively 1.56, 1.00, and 1.20, or 93, 84, and 88 chances in 100 of true changes.

The orphanage children who had two tests (Table 8) gained 0.53 sigma score. This yielded a ratio of 2.21 or 98.6 chances in 100 of a true gain in the brief interval of 0.3 month between tests. The eighteen orphanage children who had three tests gained from first to second tests almost identically with the larger group of thirty children, but lost about half of this advantage from second to third tests over an interval of 6.5 months. The chances were 96 and 80 in 100 respectively of true changes.

Intelligence Quotients

Changes in intelligence quotient tended to be more significant than changes in sigma score for the preschool group. From first to second tests there was an increase of 10.3 IQ points. This was a statistically significant increase. The preschool children who had three tests showed a mean gain of 9.4 between first and second tests and an additional gain of 2.8 points between second and third tests. The former yielded a ratio of 2.8, or 99.7 chances in 100 of a true gain between first and second tests. The change from second to third tests was insignificant, the ratio being .9 and the chances of a true gain only 82 in 100. The gain from first to third tests was statistically significant.

The twenty-two children who had four tests gained 8.6 IQ points between first and second tests, 3.5 between second and third, and 9.5 between third and fourth, making a total gain of 21.6 IQ points between first and fourth tests. The separate gains were not significant, but the gain from third to fourth tests yielded a ratio of 2.6, or 99.5 chances in 100 of a true gain. The gain from first to fourth tests was statistically significant.

The orphanage group gained 10.4 IQ points from first to second

test, with 98 chances in 100 that this was a true gain. The orphanage children who had three tests gained 9.7 points and then lost 4.5 points. The chances in 100 that these were true changes were respectively 90 and 76.

Percentiles

Changes in percentile rank followed rather closely the trend for changes in intelligence quotient, with ratios slightly lower, and the chances of true gains somewhat less. There was a significant gain in percentile rank from first to third tests and from first to fourth tests. Orphanage children gained 18.5 percentile ranks from first to second tests, with 99.7 chances in 100 that this was a true gain. There was a drop of 4.6 percentile ranks from second to third tests, but this was not a significant change.

Retests at One Week

Forty-four children were retested at one week's interval. Thirty were the orphanage children of Table 8 and the remainder were Iowa City preschool laboratory children. The IQ range on first test was from 69 to 154; the sigma score range was from —2.5 to +2.5, and the percentile range from 1 to 100. The means on each test were as follows:

Test	Mental Age	Intelligence Quotient	Sigma Score	Percentile Rank
First	44.0	99.3	— 0.14	46.0
Second	50.0	111.6	+ 0.40	62.5
Gain	6.0	12.3	0.54	16.5

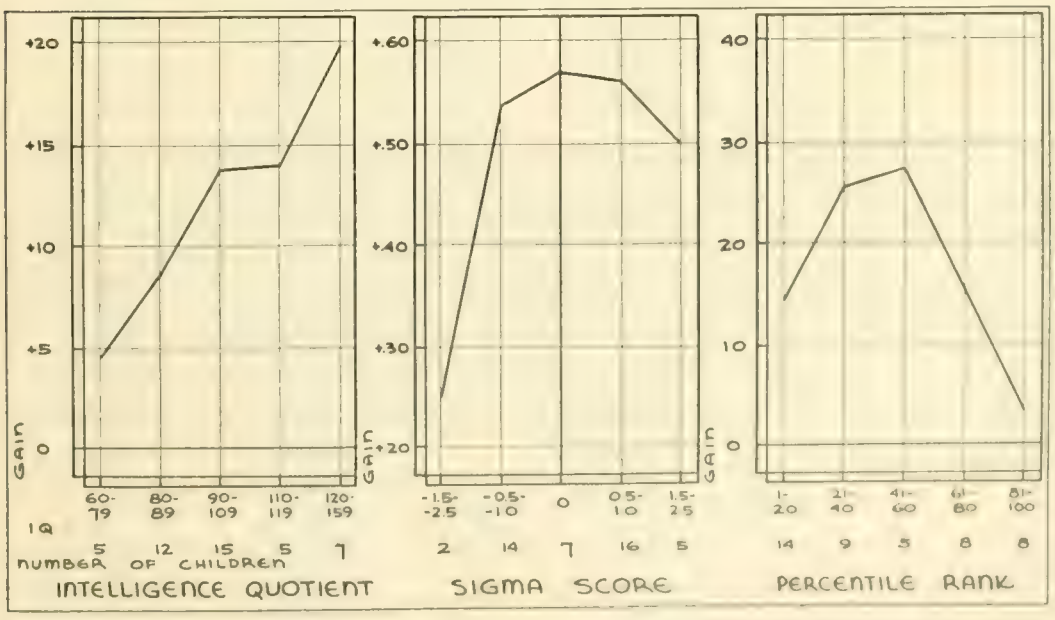


Figure 3. Gain on Retest at One Week Per Score on First Test

The effect of repetition of the test at a very brief interval was to increase the mental age by 6.0 months, the intelligence quotient by 12.3 points, the sigma score by 0.54, and the percentile rank by 16.5. In spite of the large gains, the correlations were fairly high. They were:

Mental age	.96	± .01
Sigma score	.92	± .02
Intelligence quotient	.90	± .02
Percentile rank	.89	± .02

The greatest gains were made by children who were already superior in IQ, as is shown by the following and by Figure 3:

Intelligence Quotient on First Test	Children	Gain in IQ
60 to 79	5	4.6
80 to 89	12	8.6
90 to 109	15	13.9
110 to 119	5	14.0
120 to 159	7	19.9

The amount of gain increased with successively higher initial IQ classifications, mounting to 20 points for the very superior levels. Apparently children who were superior were best able to profit by previous experience with the test. Only one child (IQ 74) lost; four others made no change, the remaining children gaining from 2 to 50 points in IQ. Eleven children gained 20 or more points.

Sigma score increases ranged from 0 to 1.5. They showed no apparent relation to the size of the initial score.

Sigma Score	Children	Gain
1.5 to 2.5	2	.25
— .5 to 1.0	14	.54
0.0	7	.57
—0.5 to —1.0	16	.56
—1.5 to —2.5	5	.50

Percentile ranks increased from 0 to 62 ranks. Children receiving very high initial ranks were limited in the gains it was possible for them to make. Consequently, the gains for the two highest groups were less than for children in the middle ranges. Children with initial ranks of 1 to 20 gained less than those with initial ranks of 21 to 40 or 41 to 60. Neither sigma score nor percentile rank showed the regularity of change with initial status that was revealed by the intelligence quotient classifications (Figure 3).

Percentile Rank	Children	Gain
1 to 20	14	14.7
21 to 40	9	25.9
41 to 60	5	27.2
61 to 80	8	15.1
81 to 100	8	3.6

Explanation of Gains

There is no reason to suppose that gains made in one week represented any valid increase in mental ability. The only plausible explanation seems to be something in the nature of practice effects from previous experience with the test. Since many of the subtests of the scale are of the form board or puzzle type and the score on them depends upon speed of completion, it does not seem unreasonable to expect higher scores on repetition of the test. After the child has once gone through the process of completing a form board, he probably can do it more efficiently and speedily a second time. A second trial at one sitting is usually accomplished in less time than the first trial. Illustrations of subtests in the Merrill-Palmer scale in which time is a factor in this way are:

1. Wallin peg board A, which consists of six round pegs to be placed in corresponding round holes. This is test item 5 in order of difficulty when done in 38 seconds or less and is then credited at the 18 to 23 months level. It is test item 15 in order of difficulty when done in 25 seconds or less and is credited at the 24 to 29 months level. It is test item 28 when done in 20 seconds or less and is credited at the 30 to 35 months level. It is test item 38 when done in 17 seconds or less and is credited at the 36 to 41 months level.
2. Wallin peg board B, which consists of six square pegs to be placed in corresponding square holes
   - Test 16, 41 seconds or less, credited at 24 to 29 months level
   - Test 26, 27 seconds or less, credited at 30 to 35 months level
   - Test 45, 23 seconds or less, credited at 36 to 41 months level
3. Sixteen cubes
   - Test 8, 13 or more in box, no time limit, 18 to 23 months level
   - Test 14, 16 in box, 125 seconds or less, 24 to 29 months level
   - Test 25, 16 in box, 100 seconds or less, 30 to 35 months level
4. Nest of cubes
   - Test 18, 250 seconds or less, 24 to 29 months level
   - Test 29, 30 seconds or less, 30 to 35 months level
5. Seguin form board, which consists of ten blocks of different forms to be placed in corresponding depressions
   - Test 24, 2 of 3 trials completed, no time limit, 30 to 35 months level
   - Test 31, 222 seconds or less, 30 to 35 months level
   - Test 36, 109 seconds or less, 36 to 41 months level
   - Test 41, 28 errors or less — total, 36 to 41 months level
   - Test 49, 72 seconds or less, 42 to 47 months level
   - Test 50, 9 errors or less — total, 42 to 47 months level
   - Test 60, 63 seconds or less, 48 to 53 months level
   - Test 77, 48 seconds or less, 54 to 59 months level
   - Test 86, 39 seconds or less, 60 to 65 months level
   - Test 93, 35 seconds or less, 66 to 71 months level
6. Three cube pyramid
   - Test 27, 17 seconds or less, 30 to 35 months level
   - Test 43, 11 seconds or less, 36 to 41 months level
   - Test 48, 9 seconds or less, 42 to 47 months level
   - Test 62, 7 seconds or less, 48 to 53 months level

Other tests similarly scored on time are two buttons, picture puzzle No. 1, little pink tower, mare and foal, four buttons, picture puzzle

No. 2, six cube pyramid, and Decroly matching game. In the total Merrill-Palmer scale there are ninety-three test items based upon thirty-nine separate tests. Fifty-eight of these test items involve fourteen separate tests which are credited at different levels on the basis of speed.

Although it seems that the gains on retests at an interval of one week must be attributed to practice effects, it does not necessarily follow that gains made over longer intervals are also entirely attributable to practice effects, although of course they may be. There is the additional possibility, which merits investigation, that such gains may be related to experiences which have brought about a real increase in intellectual ability. The data on the preschool laboratory group lend themselves to an analysis of this possibility.

The position of the orphanage children on third test (Table 8) after six months might be explained in several ways:

1. They were higher than on the first test because the practice effects from a repetition at one week's interval had not entirely worn away, even after six months, but had partially done so. Analysis of the changes made over the six-month interval reveals that the trend of changes was not a simple reversal or dropping off of the increases made at one week. The children who were below 100 IQ on their first test retained the full advantage of their increase, being 7.4 points higher on their third test, while the children who were above 100 IQ on their first test did not retain their gain, being 2.5 points inferior on their third test. These differences are also shown by correlations between initial IQ and change.

	First to Second		Second to Third		First to Third	
	Children	Change	Children	Change	Children	Change
Below 100 IQ	14	+ 8.6	11	— 0.9	14	+ 7.4
100 IQ and above	4	+ 13.5	7	— 13.0	4	— 2.5

Change from first to second test correlated positively .16 ± .14 with initial IQ, while change from first to third test correlated negatively — .60 ± .09. If by practice effects is meant memory, it remains to be explained why the brighter children did not have as serviceable a memory as the duller children.

2. A second explanation is that the gain from initial test to the third test might be accepted as a real increase in ability. However, knowledge of conditions under which the children lived would throw serious doubt on this explanation. The children lived in cottages with large numbers of other children of the same age. They had very little contact with adults and a minimum of opportunity for constructive play with materials. Most of their waking hours, with the exception of eating and sleeping, were spent in the cottages. After living under these circumstances for six months a real increase in ability could scarcely be expected.

3. A third explanation is that of regression towards the mean. This is really more of a descriptive than an explanatory term, since the phrase covers the idea of gains by children below the mean and losses by children above the mean, but does not explain why they occur. That regression towards the mean on retest is not a statistical necessity is shown by the changes from first to second tests, which cannot be accounted for by the regression phenomenon.

4. A fourth explanation is unreliability of the test. Judged by the ordinary criterion of correlation, the test was highly reliable since the correlation for retest at one week for the forty-four cases was .90 ± .02 when based on intelligence quotient. In spite of large increases, the children had not changed their respective ranks to any great extent. They had, however, shifted positions considerably over a period of six months. These shifts may be considered as due to unre-

liability, if they are the result of inequalities in standardization at the different ages. It is not definitely known that such is the case.

Comparison With Scores on Minnesota Preschool Scale

Scores on the Minnesota Preschool Scale over the same periods may help in determining the correct interpretation of the gains of the orphanage children. Twenty-seven children were given Form A of the Minnesota test either the day preceding or day following their first Merrill-Palmer test, in rotated order. Sixteen of these were orphanage children and eleven were attending the nursery school of the Iowa State Teachers' College at Cedar Falls. The orphanage children had a second Merrill-Palmer one week later. Table 9 presents the comparisons for the orphanage children.

The Minnesota Preschool Scale provides separate scoring on the basis of the verbal and nonverbal components of the test. It will be

Table 9
Comparison of Merrill-Palmer and Minnesota (Form A) Intelligence Quotients of Orphanage Children

Children	Merrill-Palmer			Minnesota Form A			
	Interval Between Tests	Test	Intelligence Quotient	Test	Intelligence Quotient		
					Total	Verbal	Nonverbal
First Minnesota at Time of First Merrill-Palmer							
16		First	96.1	First	93.3	92.4	96.0
16	1 week	Second	110.6				
First Minnesota at Time of Third Merrill-Palmer							
10		First	86.9				
10	1 week	Second	95.1				
10	6 months	Third	91.7	First	81.7	77.4	90.0
Second Minnesota at Time of Third Merrill-Palmer							
11		First	93.1	First	89.9	88.2	93.5
11	1 week	Second	110.0				
11	6 months	Third	104.7	Second	94.9	94.2	97.3

seen in Table 9 that the nonverbal intelligence quotient of the orphanage children was identical with the Merrill-Palmer intelligence quotient taken at the same time, the two means being respectively 96.1 and 96.0. On retest at one week the Merrill-Palmer intelligence quotient of these sixteen children rose 14.5 points. The verbal intelligence quotient on the Minnesota test was 3.6 points lower than the nonverbal and the two parts combined resulted in a mean within 2.8 points of the Merrill-Palmer mean. The Minnesota results were closer to the first Merrill-Palmer test than to the retest results.

Ten orphanage children received their first Minnesota test (Form A) at the time of their third Merrill-Palmer test. Their nonverbal intelli-

gence quotient was very close to the results of the third Merrill-Palmer test, the means being respectively 90.0 and 91.7. It was also close to the first Merrill-Palmer test mean of 86.9, since its value fell between the first and third tests. It was 5.1 points lower than the second Merrill-Palmer.

The relation between verbal and nonverbal Minnesota scores for the two groups is interesting. In the group of sixteen children the verbal score was 3.6 points below the nonverbal, but in the group of ten children the verbal quotient was 12.6 points inferior to the nonverbal. An hypothesis that would explain this difference is that the orphanage was notably lacking in verbal types of stimulation since the ratio of adults to child was very low in comparison with ordinary home situations. The orphanage provided only one or two adults for a cottage which consisted in some instances of as many as forty preschool-age children. The usual verbal interplay between parents and child and between older siblings and child was largely lacking. The sixteen children who showed closer correspondence between verbal and nonverbal scores were fresher from a usual home situation. For them the interval in months from admission to the orphanage to the time of the Minnesota test ranged from 2 to 31 months with a mean residence of 8.9 months. The ten children who showed the greater discrepancy between verbal and nonverbal scores had been in the orphanage 15 to 61 months, with a mean residence of 31.2 months.

The results on this second group seem to point to the conclusion that the third Merrill-Palmer test represents the status of the children in nonverbal abilities at that time, although the conclusion must be tentative in view of the small number of cases. If the change from first to third Merrill-Palmer test was due to practice effects, it is possible that these effects in some way carried over to the nonverbal portion of the Minnesota Scale.

Eleven children received a retest on the Minnesota Scale (Form A) at the same time as the third Merrill-Palmer. These eleven children were included in the group of sixteen children previously discussed. Over the interval of six months there was a gain in nonverbal quotient of 3.8 points, in verbal quotient of 6.0 points, and in total intelligence quotient of 5.0 points. It would appear that either both Minnesota and Merrill-Palmer scales show higher scores due to repetition of the test, or the children increased in ability. The change appears to be somewhat greater on Merrill-Palmer than on Minnesota retests, the

increases being 11.6 on the Merrill-Palmer, 5.0 on the total Minnesota, and 3.8 on the nonverbal Minnesota.

In the following tabulation the Cedar Falls children likewise showed close agreement between nonverbal Minnesota and Merrill-Palmer means. For them the verbal and nonverbal quotients were very similar. Five Cedar Falls children were retested on the Minnesota Form A after an interval of five months of nursery school attendance. They

	Merrill-Palmer		Minnesota Form A			
Chil-		Intelligence		Intelligence Quotient		
dren	Test	Quotient	Test	Total	Verbal	Nonverbal
11	First	122.2	First	121.8	121.6	119.3*
5			First	118.0		
5			Second (interval 5 months)	122.4		

*Based on 8 cases. The total IQ for these 8 cases was 121.9 and the verbal IQ was also 121.9.

showed an increase in total quotient of 4.4 points. The verbal and nonverbal quotients are not presented because the scores in some instances fell outside the range included in the norms.

The correlations between Merrill-Palmer quotients on first tests and Minnesota Form A quotients were:

	Children	*r*	P.E.
Total IQ	27	.84	± .04
Verbal IQ	27	.77	± .05
Nonverbal IQ	23	.77	± .06

There was no greater correspondence of the Merrill-Palmer with Minnesota nonverbal quotient than with the verbal quotient.

Twenty-one orphanage children were given Form B of the Minnesota Scale one month after Form A was administered. They were the two last groups of Table 9. The means and correlations on the two tests were:

	Intelligence Quotient			3rd Merrill-
	Total	Verbal	Nonverbal	Palmer
Form A, Mean	88.6	86.2	93.8	98.5
Form B, Mean	90.2	84.6	98.7	
Correlation	.77 ± .07	.71 ± .08	.66 ± .09	

While the means of Forms A and B were fairly close, the correlations were low for alternate equivalent forms of the same test. They were in fact lower than the correlation of Form A with the Merrill-Palmer test. The range of difference in quotients between Form A and Form B was —22 to +14 on total quotient, —22 to +23 on verbal quotient, and —33 to +7 on nonverbal quotient.

The analysis of Minnesota results on the same children has yielded nothing conclusive in regard to the interpretation of the increases on

the Merrill-Palmer retests of the orphanage children. With no evidence to the contrary, the best present judgment is to conclude that an increase of approximately five points in IQ may be expected over a six-month period from factors unknown but associated with repetition of the test at such an interval. The analyses which are to follow will show whether a similar conclusion should be drawn in regard to preschool laboratory children.

Summary

One hundred twenty-seven children were given a second Merrill-Palmer test and sixty-eight children were given a third test. Ninety preschool laboratory children received a second test, forty-seven a third, and twenty-two a fourth test. Scores on retests were higher than on the first test. Thirty orphanage children received a second test and eighteen orphanage children a third test. Scores on retests were also higher in this group.

Forty-four children received a retest at an interval of one week. Large gains on retest were found. The amount of gain was proportionate to the initial IQ level. In spite of the large gains, the correlations were fairly high, .96 ± .01 for mental age, .92 ± .02 for sigma score, .90 ± .02 for intelligence quotient, and .89 ± .02 for percentile rank. The best explanation for these gains at one week appears to be practice (memory) effects, probably due to the large weighting of scores, particularly on form boards, on the basis of speed.

Retested after six months, the orphanage children were higher than on their initial test, although lower than on the second test which was given one week after the first test. The gain over this interval was inversely proportionate to initial IQ level. No really satisfactory explanation of their position on the third test was advanced. The best present judgment appears to be to conclude that an increase of approximately five points in IQ may be expected on retest at an interval of six months from factors unknown but associated with repetition of the test at such an interval. Comparison with scores on the Minnesota Preschool Scale did not contribute to the explanation.

The explanation of gains made by the preschool laboratory children is reserved for the following section of the study on the effects of preschool attendance.

EFFECTS OF PRESCHOOL ATTENDANCE

The data on the preschool laboratory children may be analyzed in a number of ways in order to bring out the effects, if any, of preschool attendance. Several such analyses have been made:

1. Comparison of the changes between fall and spring tests with those between spring and fall tests: If preschool attendance changes intelligence, growth during periods of attendance should differ from that during comparable periods of nonattendance of the same children. This should be the case, provided that a sufficiently long period is allowed for the effects to take root. Since previous investigations (37, 38) have shown that a six-month period is sufficient to reveal significant differences on Binet tests, a similar period has been used for Merrill-Palmer analyses. Thus, gains over the winter months covering preschool attendance are compared with those over the summer vacation months of the same children.

2. Relation of changes within the six-month period of preschool attendance to actual number of days in attendance: This is a refinement of the first comparison, calling for an impact of preschool experiences cumulative in nature and becoming effective within a few days, since the maximum number of days of possible attendance in a six-month period is only 120. Such a comparison assumes that the effects of preschool attendance, if any, work with the same rapidity for different children. Also, it must be borne in mind that preschool attendance is not concentrated directed training, since preschool education has not been specifically aimed at increasing intelligence. Looking at the problem with these considerations in mind, it seems that differential changes with sheer number of days of preschool experience should not be expected. Nevertheless a positive relationship between magnitude of change and number of days in preschool has been found on Binet tests (37), warranting a similar type of comparison on Merrill-Palmer results.

3. Relation of changes to length of daily session: Some of the children in the present study attended preschool for half-day sessions and others stayed through the noonday meal, afternoon nap, and an afternoon play period. This, again, is a refined type of comparison.

4. Changes according to interval between tests: By and large, the children who had the longest intervals between tests had the greatest amount of intervening preschool attendance, since retests were given only when the children continued to be enrolled in preschool. Compari-

Table 10
Changes in Intelligence During Periods of Preschool Attendance and Nonattendance

Time of Test	Order of Test	Children	Mean Chronological Age	Sigma Score		Intelligence Quotient		Percentile Rank		Binet Intelligence Quotient		
				Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Children	Mean	Standard Deviation
Fall-Spring												
Fall	First	42	35.9	.69	.68	113.4	13.3	74.9	24.7	38	124.4	13.9
Spring	Second	42	41.5	1.19	.76	124.1	15.8	85.0	19.2	38	132.6	17.2
Difference	First and second		5.6	.50		10.7		10.1			8.2	
Spring-Fall												
Spring	First	20	38.9	.87	.75	115.2	14.9	76.9	11.4	18	128.5	21.1
Fall	Second	20	45.1	.85	.69	120.4	18.8	78.9	21.6	18	128.7	17.4
Difference	First and second		6.2	—.02		5.2		2.0			.2	
Fall-Spring-Fall												
Fall	First	19	31.4	.52	.82	107.8	12.1	65.4	27.7	17	123.6	14.5
Spring	Second	19	36.8	1.23	.77	121.5	14.0	84.5	20.3	17	136.8	15.1
Fall	Third	19	43.6	1.08	.59	124.9	13.5	87.8	12.1		13.2	
Difference	First and second		5.4	.71		13.7		19.1				
Difference	Second and third		6.8	—.15		3.4		3.3				
Spring-Fall-Spring												
Spring	First	17	38.5	.85	.74	115.7	15.2	77.3	17.8	15	130.1	21.9
Fall	Second	17	44.5	.85	.66	120.8	19.4	79.8	20.8	15	128.5	15.1
Spring	Third	17	50.6	.85	.64	124.1	14.0	84.6	16.2	15	133.1	13.9
Difference	First and second		6.0	0.0		5.1		2.5			—1.6	
Difference	Second and third		6.1	0.0		3.3		4.8			4.6	
Fall-Spring-Fall-Spring												
Fall	First	10	28.1	.80		110.4		72.0		10	125.5	
Spring	Second	10	33.3	1.45		122.5		85.6		10	138.5	
Fall	Third	10	40.4	1.25		128.2		90.3		10	134.3	
Spring	Fourth	10	45.5	1.75		142.7		98.5		10	132.0	
Difference	First and second		5.2	.65		12.1		13.6				
Difference	Second and third		7.1	—.20		6.0		4.7				
Difference	Third and fourth		5.1	.50		14.2		8.2				

Spring-Fall-Spring-Fall												
Spring	First	6	35.7	1.25		122.0		83.8		4	147.3	
Fall	Second	6	41.7	1.25		127.3		90.8		4	142.3	
Spring	Third	6	48.0	1.16		129.7		93.5		4	143.2	
Fall	Fourth	6	53.7	1.00		132.0		92.0		4	139.2	
Difference	First and second		6.0	0.0		5.3		7.0			—5.0	
Difference	Second and third		6.3	—.09		2.4		2.7			.9	
Difference	Third and fourth		5.7	—.16		2.3		—1.5			—4.0	
Combined Groups: Fall-Spring-Fall and Spring-Fall-Spring												
Fall	First and second	36	37.6	.68	.76	114.0	17.3	72.2	26.3	32	125.9	15.4
Spring	Second and third	36	43.3	1.05	.74	122.7	14.2	84.6	18.8	32	135.1	15.0
Difference	First and second and second and third		5.7	.37		8.7		12.4			9.2	
Spring	First and second	36	37.6	1.05	.78	118.8	14.9	81.1	19.5	32	133.7	18.6
Fall	Second and third	36	44.0	.97	.64	123.0	17.1	84.0	17.4	32	130.2	15.0
Difference	First and second and second and third		6.4	—.08		4.2		2.9			—3.5	
Combined Groups: Fall-Spring-Fall-Spring and Spring-Fall-Spring-Fall												
Fall	First, second, and third	26	36.0	1.07	.74	121.2	16.0	83.4	21.3	24	132.0	16.0
Spring	Second, third, and fourth	26	41.4	1.50	.64	131.9	15.4	92.4	14.7	24	136.6	16.7
Difference	First, second, and third and second, third and fourth		5.4	.43		10.7		9.0			4.6	
Spring	First, second, and third	22	37.9	1.32	.75	124.3	14.5	87.1	17.3	20	142.0	17.4
Fall	Second, third, and fourth	22	44.4	1.18	.51	129.0	12.9	90.9	9.1	20	135.2	14.9
Difference	First, second, and third and second, third and fourth		6.5	—.14		4.7		3.8			—6.8	
Combined Groups: All Fall-Spring												
Fall	All	72	39.2	.83	.68	118.1	16.4	79.0	22.6	66	127.1	15.3
Spring	All	72	44.8	1.19	.73	127.2	16.0	87.2	17.3	66	131.8	16.1
Difference	Fall and spring		5.6	.36		9.1		8.2			4.7	
Combined Groups: All Spring-Fall												
Spring	All	46	39.3	1.10	.77	120.7	16.2	82.7	19.3	43	134.9	19.0
Fall	All	46	45.7	.98	.62	124.4	16.3	84.8	17.6	43	130.4	16.0
Difference	Spring and fall		6.4	—.12		3.7		2.1			—4.5	

sons are made of the changes over half-year, year, and year and a half intervals. These comparisons are complicated by the fact that the children having the longer intervals had a greater number of Merrill-Palmer tests. Hence, if there were practice effects from repetition of the test, the comparisons here would be affected.

5. Relation of level of intelligence to the amount of preschool attendance prior to the first test: Underlying this comparison is the assumption that groups of children entering preschool in the different years in which the Merrill-Palmer tests were administered were approximately equal in ability at the time they entered. If this assumption is justifiable, the effects of preschool attendance will be revealed by comparing the test performance of groups of children who received their first test with different amounts of preschool experience to their credit.

Changes During Periods of Preschool Attendance and Nonattendance

Winter and summer changes in intelligence of preschool laboratory children are compared in Tables 10 and 11. In seventy-two cases a Merrill-Palmer test was given in the fall and a subsequent test was administered the next spring after an average interval of 5.6 months of preschool attendance. The fall tests were given in October and November and the spring tests in April and May. Of these cases, forty-two represented a first test in the fall and a second test in the spring. The remaining thirty cases represented further retests, some being additional children whose first test was administered in the spring and a few being a second set of fall-spring measurements of the same children.

Forty-six cases were given a test in the spring and a subsequent test was administered the following fall after an average interval of 6.4 months, of which about four months represented nonattendance. Of these cases, twenty represented a first test in the spring and a retest in the fall.

There was no selection as to which children received their first test at either time (fall or spring) except year of enrollment. In some years all children attending the preschools who fell within the age limits of applicability of the test were tested in the fall and again in the spring. In other years the first tests were not applied until spring.

Nineteen children were followed through three tests beginning in the fall, and ten through four tests beginning in the fall. Seventeen

Table 11

Changes During Periods of Preschool Attendance and Nonattendance: Obtained Differences and Critical Ratios

Group	Sigma Score			Percentile Rank		Intelligence Quotient		Binet Intelligence Quotient		
	Children	Obtained Difference	Difference / σDifference	Obtained Difference	Difference / σDifference	Obtained Difference	Difference / σDifference	Children	Obtained Difference	Difference / σDifference
Fall to Spring Changes										
Fall-Spring	42	.50	3.18	10.1	2.09	10.7	3.35	38	8.2	2.28
Fall-Spring-Fall	19	.71	2.75	19.1	2.42	13.7	3.22	17	13.2	2.60
Spring-Fall-Spring	17	.00	0	4.8	.75	3.3	.57	15	4.6	.87
Combined Fall-Spring-Fall and Spring-Fall-Spring	36	.37	2.09	12.4	2.30	8.7	2.52	32	9.2	2.42
Combined Fall-Spring-Fall-Spring and Spring-Fall-Spring-Fall	26	.43	2.24	9.0	1.77	10.7	2.46	24	4.6	.97
Combined all groups	72	.36	3.05	8.2	2.45	9.1	3.37	66	4.7	1.72
Spring to Fall Changes										
Spring-Fall	20	—.02	.09	2.0	.37	5.2	.97	18	.2	.03
Spring-Fall-Spring	17	.00	0	2.5	.38	5.1	.85	15	—1.6	.23
Fall-Spring-Fall	19	—.15	.67	3.3	.61	3.4	.76	17	—5.1	1.06
Combined Fall-Spring-Fall and Spring-Fall-Spring	36	—.08	.48	2.9	.66	4.2	1.11	32	—3.5	.83
Combined Fall-Spring-Fall-Spring and Spring-Fall-Spring-Fall	22	—.14	.72	3.8	.91	4.7	1.14	20	—6.8	1.32
Combined all groups	46	—.12	.82	2.1	.54	3.7	1.09	43	—4.5	1.19

were followed through three tests beginning in the spring, and six through four tests beginning in the spring.

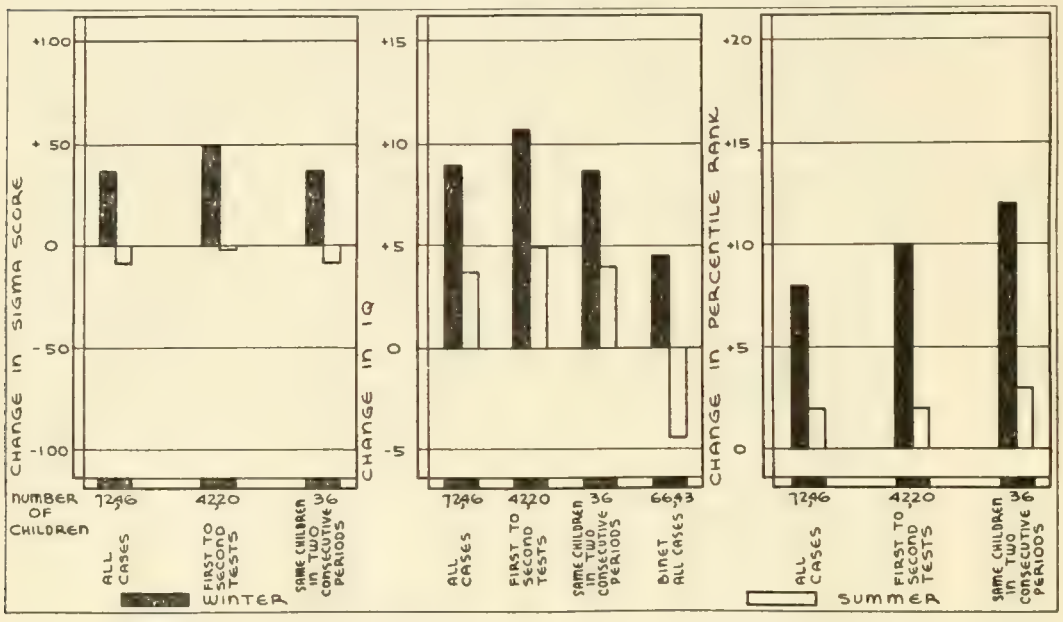


Figure 4. Changes During Six-Month Periods of Preschool Attendance and Nonattendance

Sigma Scores

The seventy-two cases of fall-spring tests had an initial mean sigma score of .83 in the fall and gained .36 sigma score during a mean period of 5.6 months of preschool attendance (Table 10 and Figure 4). The gain was statistically significant, the ratio of the difference to the standard deviation of the difference being 3.05, indicating practical certainty of a true increase. The forty-six cases of spring-fall tests had an initial mean sigma score of 1.10 in the spring and dropped —.12 sigma score over the 6.4 months preponderately of nonattendance. The loss was not a significant one, the ratio being only .82, or 79 chances in 100 of a true loss.

That these comparisons were not affected by possible inequalities in the two groups in amount of previous experience with the Merrill-Palmer test is shown by the two groups whose initial test was in each case the first Merrill-Palmer test the child had received. The forty-two children whose first test was administered in the fall had an initial mean sigma score of .69 and gained .50 sigma score in 5.6 months of preschool attendance. The twenty children whose first test was administered in the spring had an initial mean sigma score of .87 and lost .02 sigma score over the 6.2 months from spring to

fall test. The gain of the fall-spring group was significant, the ratio being 3.18; while the loss of the spring-fall group was not significant, the ratio being .09 or only 53 chances in 100 of a true loss.

There were some inequalities in initial mean scores in the above comparisons, but that such inequalities were not the cause of differences in gain is shown by comparison of the spring-fall changes of the first-to-second test group, for whom the mean was .87 and the loss .02, with the total fall-spring group, for whom the initial mean was .83 and the gain was .36.

The nineteen children having three tests beginning in the fall showed the following changes in sigma score: fall to spring .71, and spring to fall —.15. While the fall-spring change yielded a ratio of 2.75, or 99.7 chances in 100 of a true gain, the ratio of the spring-fall change was .67, or 75 chances in 100 of a true loss. The seventeen children having three tests beginning in the spring showed no change in sigma score throughout. The two groups combined yielded thirty-six cases for whom the fall-spring change could be compared with their own spring-fall change. In approximately half of the group (19 children), the fall-spring period antedated the spring-fall period, while in the remainder of the group (17 children) the spring-fall period antedated the fall-spring period, previous experience with the Merrill-Palmer test thus being equalized in the two periods. The gain of these thirty-six children from fall to spring was .37, with a ratio of 2.09, or 98 chances in 100 of a true gain. Their spring-fall change was a loss of —.08 sigma score, with a ratio of .48, or 68 chances in 100 of a true loss.

The ten children who were carried through two complete preschool years gained .65 sigma score the first preschool year, lost —.20 over the summer, and made an additional gain of .50 sigma score the second preschool year. The six children who were carried through two complete summer periods made no change the first summer, lost .09 sigma score during the preschool year, and lost an additional —.16 during the second summer. Combining the fall-spring results of these children yielded twenty-six cases who gained .43 sigma score from fall to spring, with a ratio of 2.24, or 98.6 chances in 100 of a true gain. Combining the spring-fall results of these children yielded twenty-two cases who lost —.14 sigma score from spring to fall, with a ratio of .72, or 76 chances in 100 of a true loss.

These various comparisons have shown that significant gains in

sigma score were made during periods of preschool attendance, while a tendency to slight loss over the summer period was revealed. These differences were not attributable to amount of previous experience with the test, initial status on the test, or age of the children.

Intelligence Quotients

In every instance there was some gain in IQ on repetition of the test, whether covering winter or summer periods (Tables 10 and 11 and Figure 4). However, the gains were consistently larger over the winter months than over the summer months. The seventy-two cases having a test in the fall and a retest in the spring gained 9.1 points in IQ. This gain was a significant one, the ratio of the difference to the standard deviation of the difference being 3.37. The forty-six cases having a test in the spring and a retest in the fall made a gain of 3.7 points in IQ. This gain was not significant, the ratio being 1.09, or 85 chances in 100 of a true gain.

The forty-two cases having their first Merrill-Palmer test in the fall and their second test in the spring gained 10.7 points in IQ, a significant gain, with a ratio of 3.35. The twenty cases having their first test in the spring gained 5.2 points over the summer. This difference was not significant, the ratio being .97, or 83 chances in 100 of a true gain.

The nineteen children who had three tests beginning in the fall gained 13.7 points from fall to spring and 3.4 points from spring to fall. The fall-spring gain of these children was significant, but their spring-fall gain was not, the ratio of the latter being .76, or 77 chances in 100 of a true gain. The seventeen children who had three tests beginning in the spring gained 5.1 from spring to fall and 3.3 points from fall to spring. Neither gain was significant, the ratios being respectively .85 and .57, or 80 and 72 chances in 100 of true gains. The two groups combined provided thirty-six cases for whom a fall-spring gain could be compared with a spring-fall gain, and for whom the number having their first test in the spring was approximately equal to the number having their first test in the fall. These thirty-six children gained 8.7 points in IQ over the winter months and 4.2 points over the summer months. The gain over the winter was not quite statistically significant, the ratio being 2.52 and the chances of a true gain 99.4 in 100. Their gain over the summer months yielded a ratio of 1.11, or 86 chances in 100 of a true gain.

The ten children who were followed through two complete pre-

school years beginning in the fall showed the following gains: first winter 12.1 IQ points, first summer 6.0 points, second winter 14.2 points, total gain 32.3 points. The six children who were carried through two summer periods showed the following gains: first summer 5.3 points, first winter 2.4 points, second summer 2.3 points, total gain 10.0 points. The two groups combined for fall-spring and spring-fall comparisons yielded twenty-six cases with a winter gain of 10.7 points and twenty-two cases with a summer gain of 4.7 points. The winter gains lacked complete statistical significance, the ratio being 2.46, or 99.3 chances in 100 of a true gain. The summer ratio was 1.14, or 86 chances in 100 of a true gain.

It will be seen that the gains over the winter months were usually approximately twice the gains over the summer months. In most instances the winter gains were statistically significant or approached significance, all but one ratio being 2.4 or above; in no instance was a summer gain significant, the highest ratio being 1.14. Since the number of tests and interval between tests were comparable, it seems reasonable to conclude that some condition or conditions other than practice effects, regression effects, or unreliability of the test were responsible for the differences between winter and summer changes. The only definitely known condition that varied with these periods was that the children attended preschool during the winter and not during the summer months.

Previous studies have shown that children attending preschool gain in Binet IQ over the winter months, but do not gain over the summer months (37, 38). Binet IQ's are given in Tables 10 and 11 for all children who received a Binet test (either Kuhlmann or Stanford revision) within one week of the Merrill-Palmer test. The results confirm previous findings of gains during periods of attendance and lack of gain during the summer vacation. The sixty-six children having a test in the fall and a retest in the spring made a gain of 4.7 points. This gain was not entirely significant, the ratio being 1.72, or 96 chances in 100 of a true gain. These children were very superior in initial ability, the fall Binet IQ being 127.1. The change, although not significant, was consistent in amount with previously reported findings at this level of ability. Larger gains at the average and superior levels were found in these studies. The forty-three children having a Binet test in the spring and a retest in the fall lost 4.5 points. The loss was not a significant one, the ratio being 1.19 and the chances of a true loss 87 in 100. This loss was also consistent with previous

findings on Binet retests at this level of ability over the summer months.

Thirty-eight of the forty-two children having their first Merrill-Palmer test in the fall also received a Binet test. They gained 8.2 points in Binet IQ from fall to spring. This gain yielded a ratio of 2.28, or 98.8 chances in 100 of a true gain. Eighteen of the twenty children having their first Merrill-Palmer test in the spring also had a Binet test. They did not change in Binet IQ over the summer, the mean change being .2.

Other comparisons on Binet IQ indicate gains over the winter and losses over the summer. The winter gains tended toward significance, while the summer losses yielded smaller ratios.

Correlations Between Mental Age and Chronological Age

With significant gains in IQ's over the winter months, it might be expected that the relationship between mental age and chronological age would be disturbed, with a consequently lower relationship on spring tests than on fall tests. The following correlations bear out this expectation, but the drop in relationship was small in each case. The correlation of .91 ± .01 obtained in the fall between Merrill-Palmer mental age and chronological age was reduced only to .87 ± .02 in the spring. The fall correlation of .87 ± .02 between Binet mental age and chronological age was reduced only to .83 ± .03. Thus, while significant changes in mean were made, the relationship to chronological age was only slightly affected.

	Merrill-Palmer			Binet		
	Children	Correlation	Probable Error	Children	Correlation	Probable Error
Fall	72	.91	.01	66	.87	.02
Spring	72	.87	.02	66	.83	.03

It will be noted that the initial correlation of .91 ± .01 between mental and chronological age was higher than that obtained for the total group of 281 children on first tests, for whom the correlation was .81 ± .01. The explanation for the higher relationship in this group may be that it was more homogeneous in level of ability, the children all being enrolled in preschool, while the larger group included some orphanage and other children of lower IQ, hence lower mental ages for their chronological ages.

Percentiles

The changes in percentile rank followed the trend of Merrill-Palmer IQ change, there being gains in both winter and summer periods, but

the winter gains being several times the size of the summer gains. The ratios of change in percentile rank were smaller than the ratios of change in IQ. In no case was there complete statistical significance. The winter changes ranged from 4.8 percentile ranks to 19.1 percentile ranks, with ratios varying from .75 to 2.45. The summer changes ranged from 2.0 to 3.8 percentile ranks, with ratios varying from .37 to .91. No summer change was as large as the smallest winter change.

The winter gain for the total fall-spring group of seventy-two cases was 8.2 percentile ranks, yielding a ratio of 2.45, or 99.3 chances in 100 of a true gain. The summer change for the total spring-fall group of forty-six cases was 2.1 percentile ranks, yielding a ratio of .54, or only 70 chances in 100 of a true gain.

Thus the percentile changes bear out the changes in sigma score and IQ, but the differences were less clear-cut. They indicate some basic differences between winter and summer changes.

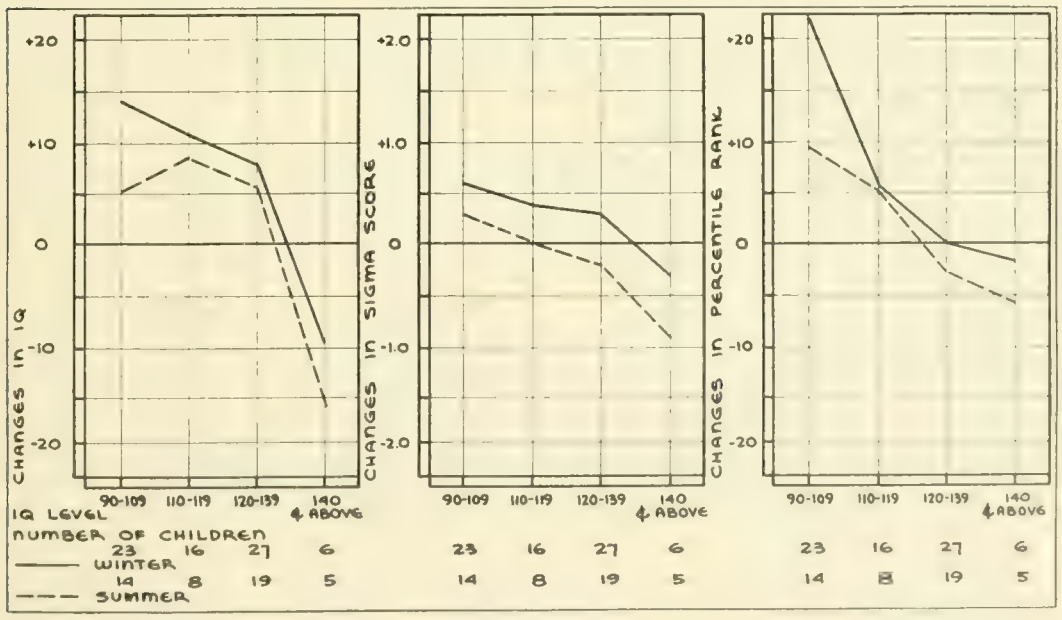


Figure 5. Changes Per IQ Level Over Periods of Preschool Attendance and Non-attendance

Comparison of Gains per IQ Level

The gains made over the winter months were inversely proportional to the intelligence level (Table 12 and Figure 5). Children who were average in intelligence (90 to 109 IQ) gained 14.0 points in IQ, those who were superior (110 to 119 IQ) gained 10.9 points, those very superior (120 to 139 IQ) gained 7.9 points, and those 140 or above in initial IQ lost 9.5 points. The correlation between initial IQ and change from fall to spring was —.43 ± .07.

Over the summer months the gain in IQ was less for each classification than over the winter months, and the loss was greater for children testing 140 or above. The trend of change in IQ appeared less regular over the summer, the change over the summer being as great for very superior children as for average children, in contrast to the difference in gain of six points between these two levels over the winter months. However, the correlation between IQ in the spring and change in IQ over the summer was —.38 ± .08, indicating essentially as high a relationship between IQ level and change over the summer as over the winter months, gains being inversely proportional to the IQ level.

Gains in sigma score were greater over the winter months than over

Table 12
Comparison of Gains in IQ Level Over Periods of Preschool Attendance and Nonattendance

IQ Level	Children	Change			Binet		Days Attendance
		IQ	Sigma Score	Percentile	Children	Change in IQ	
Fall to Spring							
90 to 109	23	14.0	.6	22.0	20	8.9	82.5
110 to 119	16	10.9	.4	5.5	15	6.5	83.9
120 to 139	27	7.9	.3	0.0	25	2.1	78.3
140 and above	6	— 9.5	—.3	— 1.8	6	—2.2	90.7
Spring to Fall							
90 to 109	14	5.1	.3	9.7	13	—3.9	
110 to 119	8	8.6	.0	5.3	8	—1.0	
120 to 139	19	5.5	—.2	— 2.7	17	—6.4	
140 and above	5	—15.4	—.9	— 5.8	5	—4.8	

the summer months for each IQ classification, and the highest level of IQ lost more in sigma score over the summer than over the winter. The trend in sigma score change over the winter months followed the trend of IQ change. Gains in percentile rank were also greater for the lower IQ levels over the winter months. The same trend was evident over the summer months, but in reduced amounts for each classification.

The Binet changes given in Table 12 are also for groups classified on Merrill-Palmer IQ. Over the winter months there was a definite trend for greater gains in Binet IQ by the children testing lowest on Merrill-Palmer. The correlation between Merrill-Palmer IQ in the fall and change in Binet IQ from fall to spring for sixty-six children was —.38 ± .07. This was nearly as high as the relationship for changes in Merrill-Palmer IQ (—.43 ± .07). Summer changes in Binet IQ were losses throughout, with no regularity of trend, the lowest and highest Merrill-Palmer groups losing equally. The correlation between

Merrill-Palmer IQ in the spring and change in Binet IQ over the summer was —.12 ± .10, indicating no tendency for summer changes in Binet IQ to be related to Merrill-Palmer level.

There was no relationship between changes in Merrill-Palmer IQ from fall to spring and initial Binet IQ, the correlation being .003 ± .08. Neither was there any relationship between summer changes in Merrill-Palmer IQ and Binet IQ in the spring, the correlation being —.04 ± .11.

Binet changes over the winter and initial Binet IQ correlated —.36 ± .07. Binet changes over the summer and Binet IQ in the spring correlated —.53 ± .08. Both correlations show that the change in Binet IQ was inversely proportional to the Binet IQ level, the tendency being somewhat more pronounced over the summer months than over the winter months.

Despite the fact that Binet changes were inversely related to initial Merrill-Palmer status and that Merrill-Palmer changes were also inversely related to initial Merrill-Palmer status, there was no relationship between Merrill-Palmer IQ change from fall to spring and Binet IQ change over the same period. The correlation was —.09 ± .08. This indicates that the same individuals were not responsible for the big changes in Merrill-Palmer that were making the big changes in Binet IQ. At the same time groups when selected on the basis of initial IQ showed a tendency towards corresponding changes in Merrill-Palmer and Binet, although different individuals within the classification were responsible for the larger gains on the two tests.

Over the summer months Merrill-Palmer changes and Binet changes showed a slight positive relationship, the correlation being .29 ± .09.

	Children	*r*
Merrill-Palmer IQ in Fall and change in Merrill-Palmer IQ Fall to Spring	72	—.43 ± .07
Merrill-Palmer IQ in Spring and change in Merrill-Palmer IQ Spring to Fall	46	—.38 ± .08
Merrill-Palmer IQ in Fall and change in Binet IQ Fall to Spring	66	—.38 ± .07
Merrill-Palmer IQ in Spring and change in Binet IQ Spring to Fall	43	—.12 ± .10
Binet IQ in Fall and change in Merrill-Palmer IQ Fall to Spring	68	.003 ± .08
Binet IQ in Spring and change in Merrill-Palmer IQ Spring to Fall	44	—.04 ± .11
Binet IQ in Fall and change in Binet IQ Fall to Spring	66	—.36 ± .07
Binet IQ in Spring and change in Binet IQ Spring to Fall	43	—.53 ± .08
Change in Merrill-Palmer and change in Binet Fall to Spring	66	—.09 ± .08

Change in Merrill-Palmer and change in Binet Spring to Fall	43	.29 ± .09
Merrill-Palmer IQ and change in Merrill-Palmer IQ on retest at one week	44	.40 ± .10

The regularity of trend in Merrill-Palmer IQ changes over the six-month periods takes on added significance when contrasted with the changes which occurred on retests at an interval of one week, previously shown in Figure 3. The contrast is revealed by Figure 6 where the two curves are brought together. The changes over one week were exactly the reverse in trend with IQ level from the changes over six months. Since the group retested at one week was a mixed one composed of preschool and orphanage children, the changes for the preschool component of the group are shown separately in the third curve of Figure 6. The contrast is even greater than when the total group

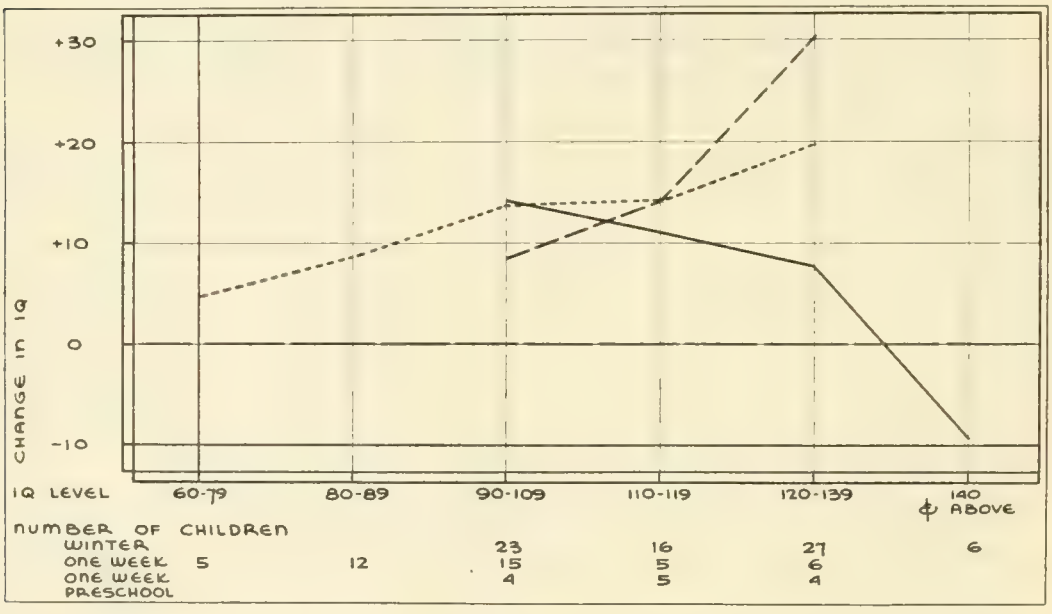


Figure 6. Changes Per IQ Level Over a Six-Month Period of Preschool Attendance Compared With Gains in One Week. (.... one week; - - - - one week preschool; — winter).

was used. The changes of preschool children per IQ level at one week and at six months, fall to spring, were:

	One Week		Six Months, Winter	
IQ Level	Children	Change	Children	Change
90 to 109	4	8.3	23	14.0
110 to 119	5	14.0	16	10.9
120 to 139	4	30.2	27	7.9
140 and above			6	— 9.5

The number of preschool cases retested after one week was very small, but the trend was regular and in the same direction as for the total group. The consistency of results permits the tentative conclusion

that the causes of change were not the same over the six-month period as over the one-week period. In regard to the retests at one week it was concluded earlier that the best explanation of change was practice effects from previous experience with the test materials, and that the children who were very superior profited most by such experience. If the hypothesis is advanced that practice effects alone constitute the correct explanation also for the changes over the six-month periods, it becomes necessary to explain why the children of average ability gained more than the children of very superior ability. On the hypothesis of practice effects alone, it would be necessary to conclude that children of average ability retain over the longer period the full advantage that is represented by quick repetition of the test, while very superior children are unable to retain their advantage in full; in fact they are unable to retain as much advantage as children of average ability. This hypothesis does not seem adequate.

A more logical hypothesis is that practice effects wore off after a time and a new set of influences operated to bring the average children closer to the level of the very superior children. Whatever these influences were, they operated both over the winter and summer periods, but in greater measure over the winter than over the summer and in greater measure for average children than for very superior children. Since the children were not in preschool during the summer, the educational influences of the preschool cannot be made to carry the full

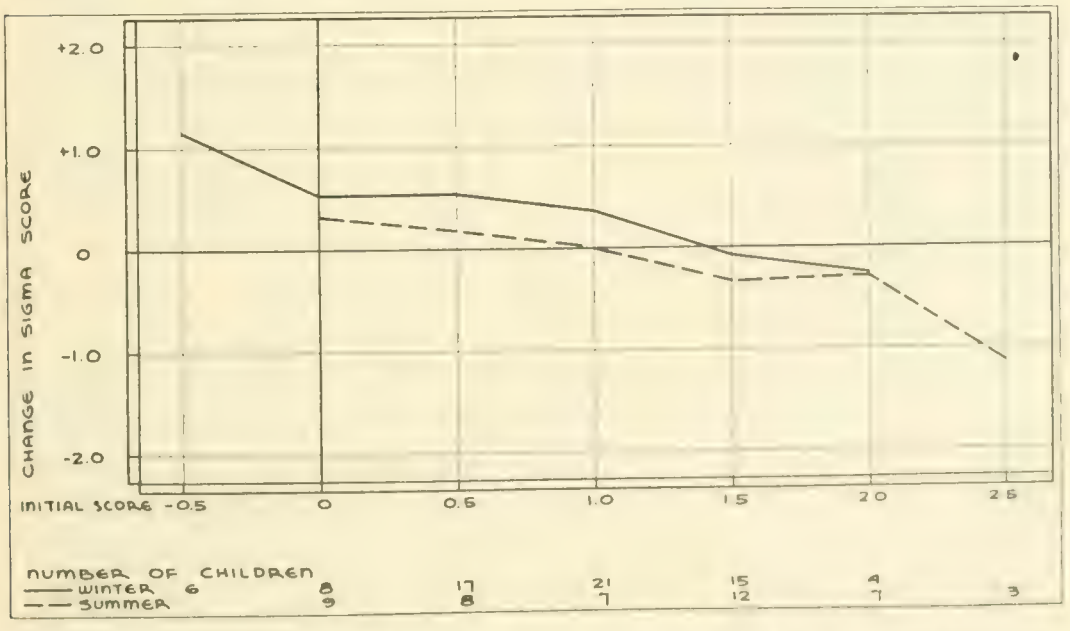


Figure 7. Changes Per Initial Sigma Score Over Periods of Preschool Attendance and Nonattendance

weight as the causative influence, although they appear to explain the difference between winter and summer gains.

Changes According to Initial Sigma Score

The sigma score changes likewise showed a regular inverse trend in magnitude with increase in initial sigma score over both winter and summer periods (Table 13 and Figure 7). Each initial sigma score classification showed a greater gain over the winter period than over the summer period. Children whose initial status was —0.5 to 1.0 sigma score gained over the winter months while those whose initial status was 1.5 or above showed losses.

The summer changes in sigma score followed the pattern of the

Table 13
Changes According to Initial Sigma Score Over Periods of Preschool Attendance and Nonattendance

Initial Sigma Score	Fall to Spring		Spring to Fall		Interval One Week	
	Children	Change	Children	Change	Children	Change
—2.5					1	1.00
—2.0					1	.00
—1.5					3	.50
—1.0					6	.66
—0.5	6	1.17			10	.50
0.0	8	.56	9	.33	8	.50
0.5	17	.53	8	.19	8	.62
1.0	21	.38	7	.00	5	.50
1.5	15	— .07	12	— .37	0	
2.0	4	— .25	7	— .29	1	.50
2.5	1	.00	3	—1.17	1	.00
Total	72	.36	46	— .12	44	.53

winter changes but the gains were smaller and the losses greater. The regularity in trend over both winter and summer periods was in contrast to the gains on retest after one week, where all initial classifications tended to gain about .5 sigma score. The mean gain for the forty-four children retested at one week was .53.

Explanation of Gains

These contrasts again point to the conclusion that the causes of change over the six-month periods differed from the causes of change in score on retest at one week. If the hypothesis is advanced that the change over the longer periods be attributed to retention of solutions from the first testing (practice effects) it is necessary to explain two facts: (1) the greater retention at each level over the winter months than over the summer months, the length of intervening time being the same, and (2) the greater retention at the average and slightly

superior levels (initial sigma scores 0 to 1.0) than at the very superior levels (initial sigma scores 1.5 to 2.5). It does not seem probable that children of high ability retain less of the advantage from previous experience with the test than children of average ability.

Another hypothesis is that commonly referred to as regression towards the mean. According to this hypothesis very superior children are expected to lose and below average children to gain, both extremes approaching the mean. This phenomenon when observed has been interpreted at times as purely statistical, the chances of loss of high-scoring children and the chances of gain of low-scoring children being automatically greater. At times the phenomenon has received a biological interpretation of the tendency of the organism to veer toward the general level of the race. There are two difficulties in acceptance of the regression-towards-the-mean hypothesis here: (1) the facts of change on retest at one week do not fit the expected trend, and (2) the winter and summer groups are selected superior groups. Instead of gaining, the children at 0.0 sigma score should not have changed, and the children at 0.5 and 1.0 sigma score should have lost.

The hypothesis which seems best to satisfy the situation is that a portion of the changes over the winter were direct responses to environmental stimulation. In view of the gains made by orphanage children over a six-month period, it seems necessary to subtract from the total change of preschool children an amount sufficient to take care of ordinary expectancy of change on retest for undetermined reasons. It is interesting to note that the change in IQ for orphanage children was roughly equivalent to the change of preschool children over the summer months. The difference between winter and summer gains may be taken as representing the amount of increase in real ability contributed by the preschool experiences.

Whether the summer gains in IQ can be given an environmental connotation remains an open question. Such an hypothesis would state that children of average ability found the environments provided by their homes over the summer months somewhat stimulating, while children of very superior ability did not find sufficient challenge in their summer environments to maintain their high level.

The hypothesis that the preschool environment has contributed to the gains is the only one which offers any explanation for the differences between winter and summer gains. We have shown previously that such differences cannot be accounted for as artifacts, since they were not due to amount of previous experience with the test (num-

ber of tests received), age, or length of interval between repetitions of the test (that is, to variations in interval between winter and summer periods).

Correlations Between Tests

The extent to which relative standing is displaced by preschool attendance is revealed by comparative correlations over winter and summer periods, given in Table 14. When the six-month period be-

Table 14
Correlations Between Tests Covering Periods of Preschool Attendance and Nonattendance

Score	Fall to Spring			Spring to Fall		
	Children	r	P.E.	Children	r	P.E.
Sigma Score	72	.48	± .06	46	.67	± .06
Percentile Rank	72	.54	± .06	46	.59	± .06
Intelligence Quotient	72	.60	± .05	46	.73	± .05
Binet Intelligence Quotient	66	.61	± .05	43	.65	± .06

tween tests covered preschool attendance, the correlation was reduced as compared with a six-month period preponderantly of nonattendance. From spring to fall the correlation between sigma scores was .67 ± .06, while from fall to spring the correlation was .48 ± .06. Similar reductions in correlations of percentile rank from .59 ± .06 to .54 ± .06 and of intelligence quotient from .73 ± .05 to .60 ± .05 are shown. It should be kept in mind that the children in the spring-to-fall group had had at least six months of preschool attendance before their spring test, and that their relative standing on the spring test was affected by such preschool experiences. The correlation, however, represents the consistency with which their placement on the spring test was maintained through the vacation months.

The correlation between Merrill-Palmer intelligence quotients was practically identical with the correlation between Binet intelligence quotients from fall to spring (.60 and .61), but there was possibly a somewhat greater consistency in Merrill-Palmer intelligence quotients over the summer period than in Binet intelligence quotients over the summer, represented by the correlations of .73 ± .05 and .65 ± .06 respectively.

Age Differences

Changes over the winter months by age groups are compared in Table 15. There were twelve two-year-old children, thirty three-year-

Table 15
Changes in Intelligence During Periods of Preschool Attendance and Nonattendance: Age Differences

Age, Months	Children	First Mean	Standard Deviation	Second Mean	Standard Deviation	Difference
Fall to Spring Changes						
Sigma Score						
18 to 29	12	0.6		1.1		0.5
30 to 41	30	0.8	0.7	1.3	0.7	0.5
42 to 53	30	0.9	0.5	1.2	0.7	0.3
Intelligence Quotient						
18 to 29	12	106.3		116.2		9.9
30 to 41	30	115.4	14.8	126.5	16.3	11.1
42 to 53	30	125.6	16.0	132.4	14.1	6.8
Percentile Rank						
18 to 29	12	63.9		80.0		16.1
30 to 41	30	76.6	23.8	86.9	18.0	10.3
42 to 53	30	87.6	12.9	90.3	11.6	2.7
Spring to Fall Changes						
Sigma Score						
30 to 41	24	1.2	0.8	1.0	0.6	— 0.2
42 to 53	19	1.0	0.7	0.9	0.6	— 0.1
Intelligence Quotient						
30 to 41	24	119.0	15.5	121.5	13.9	2.5
42 to 53	19	122.9	18.3	127.9	20.0	5.0
Percentile Rank						
30 to 41	24	81.7	20.0	83.4	18.6	1.7
42 to 53	19	83.1	19.3	84.6	16.9	1.5

old children, and thirty four-year-old children who received a test in the fall and again the following spring. The three age groups were unequal in initial status, the mean increasing with age. By each method of scoring, the four-year-old group, which had the highest initial status, made the smallest gain.

Each age group showed less change over the summer months than over the winter months. There were twenty-four three-year-old children who received a test in the spring and another test the following fall. They lost 0.2 points in sigma score as compared with a gain of 0.5 over the winter months by three-year-old children. They gained 2.5 points in IQ, while the winter gain at this age was 11.1 points. Their gain in percentile rank was 1.7 over the summer, while the winter gain was 10.3 ranks. The nineteen four-year-old children who were tested spring and fall also showed somewhat less change than children of the same age tested fall and spring, although the differences were very small at this age.

Sex Differences

There were thirty-nine boys and thirty-three girls who received a

test in the fall and another test in the spring. By each method of scoring the boys were lower in initial score and made greater gains, but their added gain was not sufficient to bring them to an equal basis with the girls in the spring (Table 16). The gains for the boys were .41 sigma score, 9.7 points IQ and 9.4 percentile ranks. These gains yielded ratios of 2.6, 2.8, and 2.0, or 99.5, 99.7, and 98 chances in 100 of true gains by the boys over periods of preschool attendance. The gains of the girls were .32 sigma score, 8.3 points IQ and 6.6 percentile ranks. The ratios were 1.9, 2.0, and 1.4, and the chances in 100 of true gains were 97, 98, and 92 respectively.

There were twenty-seven boys and nineteen girls who received a

Table 16
Changes in Intelligence During Periods of Preschool Attendance and Non-Attendance: Sex Differences

Time of Test	Children	Chronological Age, Mean	Sigma Score		Intelligence Quotient		Percentile Rank	
			Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Boys								
Fall	39	38.6	.73	.65	115.6	14.8	76.2	24.1
Spring	39	44.2	1.14	.77	125.3	15.6	85.6	17.3
Difference		5.6	.41		9.7		9.4	
Ratio			2.6		2.8		2.0	
Girls								
Fall	33	39.9	.94	.70	121.2	17.1	82.4	20.0
Spring	33	45.6	1.26	.68	129.5	15.6	89.0	17.1
Difference		5.7	.32		8.3		6.6	
Ratio			1.9		2.0		1.4	
Boys								
Spring	27	38.4	.89	.74	115.4	13.0	77.8	20.4
Fall	27	44.5	.89	.67	120.4	15.7	80.9	21.0
Difference		6.1	0		5.0		3.1	
Ratio					1.3		.6	
Girls								
Spring	19	40.7	1.39	.72	128.3	17.6	89.6	15.8
Fall	19	47.0	1.10	.50	130.1	16.2	90.3	8.6
Difference		6.3	— .29		1.8		.7	
Ratio			1.5		.33		.2	

test in the spring and another test the following fall. The boys made no change in sigma score over this period of nonattendance, but did gain 5.0 points in IQ. The ratio of 1.3 indicated 90 chances in 100 of a true gain in IQ. The gain of 3.1 percentile ranks yielded a ratio of .6, or only 73 chances in 100 of a true gain. The girls started at a higher level and lost .29 sigma score, with a ratio of 1.5, or 93 chances in 100 of a true loss over the period of nonattendance. They changed

only 1.8 IQ points and .7 percentile rank, neither gain approaching significance.

Changes According to Number of Days in Attendance Between Tests

Changes according to the number of days of preschool attendance between tests are presented in Table 17. Inspection of the changes shows no regular trend with increasing amount of attendance by ten day classifications. In fact, the thirty-two children who attended 40 to 79 days made greater gains by each method of scoring of the Merrill-Palmer data than the forty children who attended 80 to 119 days. On Binet tests the reverse was true, the children having the larger amounts of attendance making the greater gains. The thirty-one children attending 40 to 79 days gained only 1.6 points in Binet IQ, while the thirty-five children attending 80 to 119 days gained 7.6 points. The Binet means in the fall were equal, differing only 0.9 point, while the spring mean of the longer attendance group was 6.9 points higher than that of the shorter attendance group. This difference in Binet IQ with increased amount of attendance corresponds with previous findings (37, 38). Apparently, attendance does not affect the Merrill-Palmer results as it does the Binet.

The correlations between the number of days in attendance between fall and spring tests and the magnitude of change in Merrill-Palmer scores were slightly negative, the correlations for the seventy-two children being as follows:

	r	P.E.
Sigma Score	—.12	± .08
Intelligence Quotient	—.16	± .08
Percentile Rank	—.02	± .08
Binet Intelligence Quotient	.25	± .08

The Binet correlation was lower than the .54 ± .04 found previously (37) for Iowa City children between length of attendance and percentiles of IQ for 120 children on first to second tests.

In interpreting the above correlations, it should be kept in mind that children of various levels of initial ability were included. Previous analyses have shown that the magnitude of change is inversely related to the initial intelligence quotient level. Therefore, children who are of different levels of ability who attend the same number of days may be expected to make different amounts of gain. Thus there is still a possibility of a positive relationship between the number of days in attendance and magnitude of change in intelligence quotient when the

Table 17
Changes According to Number of Days of Preschool Attendance Between Tests

Days Attendance	Children	Chronological Age, Months, Fall	Mean Sigma Score			Mean Intelligence Quotient			Mean Percentile Rank			Mean Binet Intelligence Quotient			
			Fall	Spring	Change	Fall	Spring	Change	Fall	Spring	Change	Children	Fall	Spring	Change
40 to 49	3	44.3	1.1	1.6	0.5	128.6	144.6	16.0	94.3	99.3	5.0	3	140.3	132.7	— 7.6
50 to 59	4	38.5	1.0	1.3	0.3	120.7	130.7	10.0	85.2	95.7	10.5	4	126.8	131.5	4.7
60 to 69	11	33.8	0.6	1.2	0.6	109.3	123.2	13.9	66.8	83.5	16.7	11	123.3	125.8	2.5
70 to 79	14	36.5	0.8	1.2	0.4	116.9	124.3	7.4	78.6	85.2	6.6	13	126.2	128.2	2.0
80 to 89	12	42.8	1.0	1.3	0.3	122.3	133.2	10.9	83.3	90.7	7.4	11	131.5	136.5	5.0
90 to 99	15	40.1	0.8	1.1	0.3	118.0	123.8	5.8	79.8	83.7	3.9	14	125.9	135.9	10.0
100 to 109	10	42.1	0.8	1.1	0.3	116.7	128.5	11.8	81.1	87.6	6.5	8	124.0	129.6	5.6
110 to 119	3	39.3	0.8	0.8	0.0	131.6	122.6	— 9.0	74.6	88.6	14.0	2	131.0	143.5	12.5
40 to 79	32	36.5	0.8	1.3	0.5	115.8	126.6	10.8	76.9	87.3	10.4	31	126.6	128.2	1.6
80 to 119	40	41.1	0.9	1.1	0.2	120.0	127.7	7.7	80.8	87.1	6.3	35	127.5	135.1	7.6

group is made homogeneous with respect to initial ability. This supposition may be tested in two ways: by partial correlation keeping initial level constant and by zero-order correlations within groups selected on the basis of initial ability.

Measures Correlated	Children	r	P.E.
Initial IQ with change in IQ	72	—.43	± .07
Initial IQ with number of days of preschool attendance	72	.03	± .08
Change in IQ with number of days of preschool attendance:			
When initial IQ is held constant by partial correlation	72	—.16	± .08
For initial IQ's 90 to 119	38	.05	± .11
For initial IQ's 120 and above	34	—.35	± .10

The inverse relationship between initial intelligence quotient and magnitude of change shown in previous analyses is represented here by the negative correlation of —.43 ± .07. Initial IQ level was not related to the number of days in attendance during the year, as is shown by the correlation of .03 ± .08. The fact that there was an inverse relationship between initial IQ and gain did not materially affect the correlation between change in IQ and number of days in attendance, however, since holding initial IQ constant by means of partial correlation resulted in a relationship of —.16 ± .08.

Dividing the children into two groups according to initial IQ and computing the correlations for each group between number of days in attendance and magnitude of change did not produce a greater positive relationship. In fact the very superior group (IQ's 120 and above on initial test) showed a negative correlation of —.35 ± .10. Thus there was a tendency for the children within the very superior group who attended the greater number of days to make less gain than the children who attended a smaller amount. Within the 90 to 119 IQ group, the relationship was practically zero, revealing no tendency for the children attending the greater number of days to make gains different from the gains made by the children who attended fewer days.

Changes With Length of Daily Session

Of the seventy-two children having tests in the fall and again in the spring, fifty-four attended half-day sessions only, sixteen attended all-day sessions from early morning to late afternoon including midday meal and afternoon nap, and two children transferred from one session to the other. In a previous study of Binet changes (37) some differences in gains were found with length of session. The analyses of Binet IQ changes here bear out the previous findings in showing greater gains for the children attending the whole-day sessions, the

gain for half-day children being 4.0 points in IQ and for full-day children 8.3 points. The reverse was true for Merrill-Palmer results, however, the children attending the half-day sessions making the greater gains. The gain for half-day children was 0.4 in sigma score as compared with 0.1 for the all-day group, 9.5 IQ points as compared with 5.7 points, and 8.9 percentile ranks as compared with 2.3 percentile ranks. Why there was this reversal of effects in Merrill-Palmer and Binet results is difficult to explain. The findings

Table 18
Changes with Length of Daily Session

Time of Test	Children	Chronological Age, Months	Mean Sigma Score	Mean Intelligence Quotient	Mean Percentile Rank	Binet	
						Children	Mean Intelligence Quotient
All Day							
Fall	16	40.4	1.0	121.5	83.3	14	129.9
Spring	16		1.1	127.2	85.6	14	138.2
Difference			0.1	5.7	2.3		8.3
Half Day							
Fall	54	38.9	0.8	117.4	78.4	50	126.7
Spring	54		1.2	126.9	87.3	50	130.7
Difference			0.4	9.5	8.9		4.0

are in line with other analyses in showing more clear-cut relationship of preschool attendance with Binet changes than with Merrill-Palmer changes.

The number of cases in the all-day group was small, but the Binet results corresponded closely to previous results in another study on forty-seven pairs of all-day and half-day children matched on IQ and age (37). The forty-seven all-day children gained 10.5 points in IQ (standard deviation 15.4) and the forty-seven half-day children gained 4.7 points in IQ (standard deviation 12.4). The difference divided by the standard deviation of the difference yielded a ratio of 2.1, or 98 chances in 100 of a true difference in favor of greater gain by the all-day children.

Changes According to Interval Between Tests

Tables 19 and 20 compare the changes over different intervals between tests: one-half year (5 to 7 months), one year (11 to 13 months), and one and one-half years (15 to 20 months).

Sigma Scores

There was no consistent trend in gain in sigma score with increasing interval between tests. The 127 children having 5 to 7 months be-

Table 19

Changes According to Interval Between Tests

Tests		Children	Mean Chronological Age	Sigma Score		Intelligence Quotient		Percentile Rank		Binet Intelligence Quotient		
				Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Children	Mean	Standard Deviation
5 to 7 Months Interval												
Earlier to later	Earlier	127	38.4	.93	.76	118.4	16.3	80.0	22.1	119	129.7	17.5
	Later	127	44.3	1.12	.70	125.6	15.8	86.2	17.6	119	131.6	16.7
First to second	First	68	36.5	.77	.78	114.3	15.0	75.7	23.9	63	125.3	16.8
	Second	68	42.2	1.12	.78	123.4	17.1	83.1	20.7	63	131.0	17.5
Second to third	Second	41	40.2	1.08	.76	121.7	17.3	82.3	21.2	38	133.8	16.2
	Third	41	46.5	.99	.60	124.6	13.1	87.1	13.4	38	132.6	14.1
Third to fourth	Third	18	41.3	1.19	.56	126.2	13.4	91.0	8.3	18	136.2	18.2
	Fourth	18	46.8	1.41	.15	136.2	11.5	95.7	5.0	18	131.2	16.1
11 to 13 Months Interval												
Earlier to later	Earlier	53	35.3	.96	.82	116.2	15.0	78.0	22.8	51	130.2	18.1
	Later	53	47.3	1.11	.64	128.4	15.2	88.6	13.7	51	132.6	16.2
First to third	First	36	35.1	.74	.78	112.6	14.2	73.0	23.9	35	128.3	19.2
	Third	36	47.2	1.00	.62	125.1	13.7	86.8	14.0	35	133.5	15.4
Second to fourth	Second	14	36.4	1.46		125.1		88.4		13	138.2	13.2
	Fourth	14	48.3	1.50		139.6		96.2		13	134.2	17.7
15 to 20 Months Interval												
Earlier to later	Earlier	24	31.1	.88	.92	111.8	15.3	72.6	26.5	24	127.2	19.5
	Later	24	48.5	1.22	.76	132.9	16.6	90.8	18.2	24	126.4	18.1
First to fourth	First	18	30.6	1.00		114.9		77.7		18	132.3	19.4
	Fourth	18	48.0	1.47		138.2		95.9		18	131.8	16.0

Table 20
Changes According to Interval Between Tests: Obtained Differences and Critical Ratios*

Order of Test	Children	Chronological Age, Difference	Sigma Score		Intelligence Quotient		Percentile Rank		Days Attendance Between Tests
			Difference	Ratio	Difference	Ratio	Difference	Ratio	
5 to 7 Months Interval									
Earlier to later	127	5.9	.19	2.1	7.2	3.6	6.2	2.5	70.8
First to second	68	5.7	.35	2.6	9.1	3.3	7.4	1.9	70.8
Second to third	41	6.3	—.09	.6	2.9	.9	4.8	1.2	75.7
Third to fourth	18	5.5	.22		10.0		4.7		59.6
11 to 13 Months Interval									
Earlier to later	53	12.0	.15	1.1	12.2	4.2	10.6	2.7	139.9
First to third	36	12.1	.26	.8	12.5	3.8	13.8	3.0	142.2
Second to fourth	14	11.9	.04		14.5		7.8		133.1
15 to 20 Months Interval									
Earlier to later	24	17.4	.34	1.4	21.1	4.6	18.2	2.7	200.9
First to fourth	18	17.4	.47		23.3		18.2		200.5

*Difference divided by standard error of difference.

tween tests gained .19, the fifty-three children having 11 to 13 months between tests gained .15, and the twenty-four children having 15 to 20 months between tests gained .34. There was a slight decrease in the correlation between tests with increasing length of interval, the correlations (Table 21) being .54, .50, and .47 respectively, all being low. In interpreting these low relationships, it should be kept in mind that retests at a one-week interval yielded a correlation of .92 ± .02 and that retests over 5 to 7 months yielded a higher relationship when the period intervening was summertime (.67 ± .06) than when it was wintertime and the children were attending preschool (.48 ± .06). The low relationships above thus are not attributable entirely to unreliability of the test. There is the possibility that some of the decrease in consistency with lengthening interval from one week to six months is caused by uneven standardization at different ages, but that does not account for the difference between winter and summer periods.

Intelligence Quotients

There was a greater amount of gain in IQ with longer interval between tests. The 127 children with 5 to 7 months between tests gained 7.2 points, the fifty-three children having 11 to 13 months between tests gained 12.2 points, and the twenty-four children having 15 to 20 months between tests gained 21.1 points. Thus the children having 11 to 13 months between tests gained 5.4 points more than those having 5 to 7 months between tests, and those having 15 to 20

Table 21
Correlations According to Interval Between Tests

Order of Test	Children	Sigma Score		Intelligence Quotient		Percentile Rank	
		r	P.E.	r	P.E.	r	P.E.
5 to 7 Months Interval							
Earlier to later	127	.54	± .05	.63	± .04	.55	± .04
First to second	68	.55	± .06	.67	± .05	.51	± .06
Second to third	41	.58	± .07	.57	± .07	.61	± .07
11 to 13 Months Interval							
Earlier to later	53	.50	± .07	.52	± .07	.26	± .09
First to third	36	.42	± .09	.50	± .08	.32	± .10
15 to 20 Months Interval							
Earlier to later	24	.47	± .11	.50	± .11	.49	± .11

months between tests gained 7.8 points more than those having 11 to 13 months and 13.2 points more than those having only 5 to 7 months between tests. The gain over each interval was statistically significant, the ratios being 3.6, 4.2, and 4.6 respectively.

In the 5 to 7 months' interval there was no intervening tests between the initial and end measurement. In the longer intervals, there were intervening tests. In the 11 to 13 months' interval, the majority of the children had one intervening test, and in the 15 to 20 months' interval, the majority had two intervening tests.

In the shortest interval (5 to 7 months) the gain was approximately equal between first and second and third and fourth tests, but there was a reduction in gain from second to third test. Since the amount of preschool attendance was as great in this period as in the other periods, the difference cannot be accounted for on this basis. The cause, or causes, of the difference are not apparent.

The correlations between retests decreased with increasing length of interval between tests (Table 21), the relationships being respectively .63, .52, and .50. These relationships are slightly higher than those for sigma scores, but are also low. They can be compared with the correlation at one-week interval of .90 ± .02. It should be remembered that when the period intervening was summertime the correlation at six-months interval was .73 ± .05, as compared with .60 ± .05 when the intervening period was wintertime and the children attending preschool. These comparisons suggest that the cause of the low relationship is not entirely a faulty test, but that children have actually grown at different rates. It is possible that uneven standardization at the different ages may have added to the decrease in relationship over the longer intervals, but it is not the sole factor.

Binet intelligence changes showed no appreciable change corresponding to length of interval between tests. The 119 children with 5 to 7 months between tests gained 1.9 points, the fifty-one children with 11 to 13 months between tests gained 2.4 points, and the twenty-four children with 15 to 20 months between tests lost 0.8 points. There was a gain of 5.7 points over a six-month interval when the tests considered were first to second tests, but there were losses of —1.2 points and —5.0 points over a similar interval when the tests were second to third and third to fourth tests. The groups were all very superior on initial tests, the means varying from 125.3 to 138.2.

Percentiles

There was an increase in the amount gained in percentile rank with increasing length of interval between tests. The gain over 5 to 7 months was 6.2 percentiles, over 11 to 13 months 10.6, and over 15 to 20 months 18.2 percentiles. While not entirely statistically significant, the ratios of the differences to the standard deviation of the differences were 2.5, 2.7, and 2.7 respectively, or 99.4, 99.7, and 99.7 chances in 100 of true gains. There was a tendency for the gains to be greater between first tests and later tests than when the interval was between two tests both of which were later in the series. This difference seems to be largely due to the fact that the children were approaching maximum scores and the possibilities of gain were more restricted when the initial status was higher.

The correlations showed some peculiarities of trend, none being high. They were respectively: 5 to 7 months .55, 11 to 13 months .26, and 15 to 20 months .49. The relationship at one week interval was .89 ± .02, at six months over the summer .59 ± .06 and six months over the winter covering preschool attendance .54 ± .06. The relationships tended to be lower than for intelligence quotients.

Age Differences

There was a decrease in the amount of gain with increase in age from two to four years (Table 22) by each method of scoring. There was greater gain at each age when the change was from first to second or first to third tests than when it was between two tests later in the series. At each age there was a greater gain in IQ over the longer interval (11 to 13 months) than over the shorter interval (5 to 7 months). Percentile ranks showed the same trend with one exception, but this trend was not true of sigma score changes.

Table 22
Changes According to Interval Between Tests: Age Differences

Age, Months	Children	Sigma Score					Intelligence Quotient					Percentile Rank				
		Mean	Standard Deviation	Mean	Standard Deviation	Difference	Mean	Standard Deviation	Mean	Standard Deviation	Difference	Mean	Standard Deviation	Mean	Standard Deviation	Difference
5 to 7 Months Interval																
Earlier to Later Test																
18 to 29	17	.76		1.29		.53	108.9		118.8		9.9	70.5		84.6		14.1
30 to 41	64	.97	.84	1.16	.70	.19	117.1	15.9	124.4	15.1	7.3	78.8	29.2	85.4	16.1	6.6
42 to 53	45	.94	.58	1.00	.64	.06	123.5	16.8	129.7	16.7	6.2	85.0	16.1	87.6	14.6	2.6
First to Second Test																
18 to 29	17	.76		1.29		.53	108.9		118.8		9.9	70.5		84.6		14.1
30 to 41	28	.71	.84	1.05	.80	.34	112.5	15.8	120.6	16.3	8.1	73.3	24.7	79.4	29.1	6.1
42 to 53	23	.85	.58	1.09	.68	.24	120.5	14.7	130.1	18.9	9.6	82.5	18.3	86.6	16.2	4.1
11 to 13 Months Interval																
Earlier to Later Test																
18 to 29	12	.79		1.00		.21	108.6		121.8		13.2	69.8		83.0		13.2
30 to 41	27	1.12	.86	1.33	.58	.21	119.3	16.2	132.8	14.8	13.5	81.0	21.1	92.9	11.2	11.9
42 to 53	14	.78		.78		.00	116.8		125.6		8.8	79.4		85.2		5.8
First to Third Test																
18 to 29	10	.65		1.05		.40	107.1		123.7		16.6	65.9		85.5		19.6
30 to 41	16	.84		1.12		.28	114.9		126.7		11.8	75.8		89.8		13.0
42 to 53	10	.65		.75		.10	114.2		124.0		9.8	75.7		83.4		7.7

Table 23

Changes According to Interval Between Tests: Sex Differences

Sex	Children	Sigma Score					Intelligence Quotient					Percentile Rank				
		Mean	Standard Deviation	Mean	Standard Deviation	Difference	Mean	Standard Deviation	Mean	Standard Deviation	Difference	Mean	Standard Deviation	Mean	Standard Deviation	Difference
5 to 7 Months Interval: Earlier to Later Test																
Boys	70	.78	.68	1.00	.70	.22	114.4	13.5	122.1	14.8	7.7	76.3	22.4	83.5	18.6	7.2
Girls	57	1.12	.81	1.27	.66	.15	123.2	17.9	129.9	15.8	6.7	84.5	20.9	89.5	15.8	5.0
Difference		.34		.27			8.8		7.8			8.2		6.0		
5 to 7 Months Interval: First to Second Test																
Boys	35	.64	.64	1.00	.81	.36	110.9	11.2	119.7	15.7	8.8	72.3	23.3	80.2	21.4	7.9
Girls	33	.91	.88	1.26	.72	.35	117.9	17.4	127.2	17.5	9.3	79.3	23.9	86.3	19.5	7.0
Difference		.27		.26			7.0		7.5			7.0		6.1		
11 to 13 Months Interval: Earlier to Later Test																
Boys	31	.74	.68	.95	.64	.21	112.2	13.5	124.5	15.6	12.3	73.0	23.7	84.8	16.3	11.8
Girls	22	1.27	.90	1.34	.59	.07	122.0	14.9	134.0	12.7	12.0	85.1	19.4	94.0	5.5	8.9
Difference		.53		.39			9.8		9.5			12.1		9.2		

Sex Differences

The girls had a somewhat higher initial score than the boys (Table 23), and the boys tended to gain slightly more than the girls, although their advantage in gain was not sufficient to bring them to an equal level with the girls. The differences in gains were very small.

Relation to Amount of Attendance Prior to First Test

Since the Merrill-Palmer testing program was begun in the spring of the year, the children tested at that time had had at least six months' experience in attending preschool before receiving their first Merrill-Palmer test. They had, however, received a Binet test shortly after entrance to preschool, and they received a Binet retest within one week of the Merrill-Palmer test, in rotated order. In subsequent years the children were given a Merrill-Palmer test a few weeks after entrance to preschool and a Binet test within one week of the Merrill-Palmer, in rotated order. This procedure resulted in a group of 111 children who had their first Merrill-Palmer test within 0 to 3 months after entrance to preschool (on the average 34.7 calendar days, the mean attendance being 21.3 days) and another group of 64 children who had their first Merrill-Palmer test after they had been in preschool 6 to 11 months. An additional group of 11 children had attended 16 or more months prior to their first Merrill-Palmer test (Table 24 and Figure 8).

If preschool attendance increases intelligence as measured by the

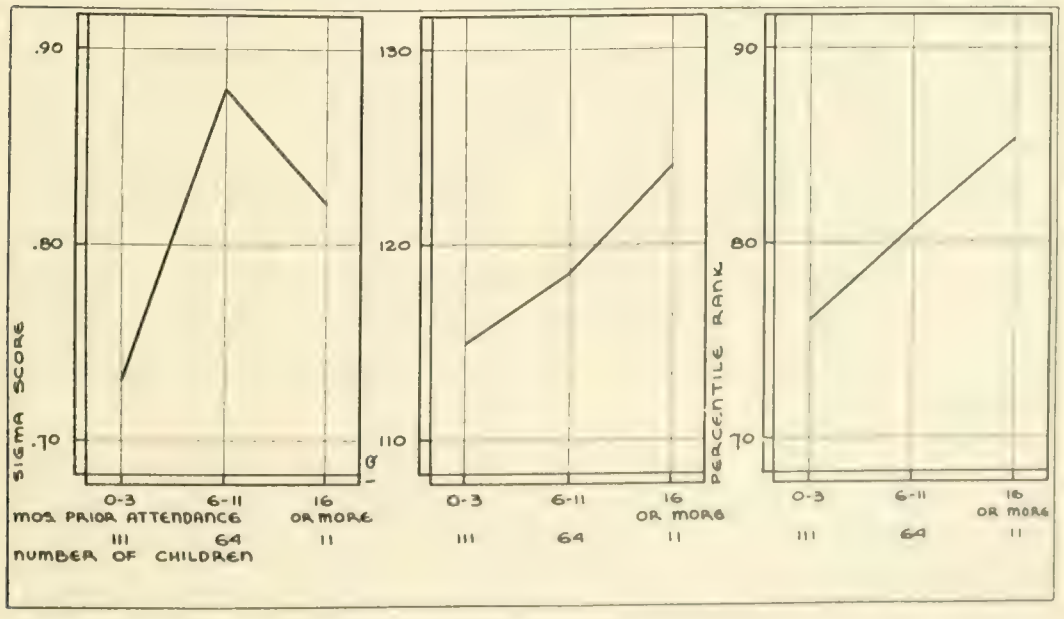


Figure 8. Relation to Amount of Preschool Attendance Prior to First Test

Merrill-Palmer scale, it is reasonable to expect scores made by the children having the greater amount of attendance prior to their test to be higher, providing no selective factors are operating. Since the basis of selection was the year the child was enrolled and the number of cases was fairly large, the groups should be expected to be approximately equal at the beginning of their preschool career. Hence differences, if found, may be considered related to the differences in preschool experience.

Sigma Scores

From Table 24 and Figure 8 it will be seen that the mean sigma score of the 64 children who received their first test when they had been in preschool 6 to 11 months was higher than the mean sigma score of the 111 children who were tested shortly after entrance to preschool. The means were respectively .88 and .73. The difference of .15 sigma score yielded a ratio of 1.31, with 90 chances in 100 of a true superiority of the longer attendance group.

The eleven children who had attended sixteen months or more before their first Merrill-Palmer test received a mean sigma score .09 higher than the new entrants and .06 lower than the group attending 6 to 11 months.

Intelligence Quotients

The children having longer periods of preschool attendance prior to their first Merrill-Palmer test in each instance showed higher mean intelligence quotients than the children who had shorter periods of attendance prior to their first test. The 111 children tested near the time of entrance to preschool had a mean IQ of 114.8, with a standard deviation of 16.1. The sixty-four children first tested after 6 to 11 months of preschool had a mean 3.6 points higher than this group and the eleven children first tested after sixteen or more months' attendance were 9.2 points higher in IQ than the new entrants. The difference between the 6 to 11 months group and the new entrants yielded a ratio of 1.52, or 93 chances in 100 of a true superiority of the longer attendance group.

The Binet intelligence quotients were not superior for the children who had been in attendance 6 to 11 months, the difference between them and the new entrants being only 1.7 points. This at first seems somewhat surprising in view of the fact that increases in Binet IQ have been found associated with preschool attendance. That the 6 to

Table 24

Relation to Amount of Preschool Attendance Prior to First Test

	Months Prior Attendance			Difference			Ratio
	0 to 3	6 to 11	16 or More	0 to 3 and 6 to 11	6 to 11 and 16 or More	0 to 3 and 16 or More	0 to 3 and 6 to 11
Children	111	64	11				
Days prior attendance	21.3	112.1	285.0	90.8	172.9	263.7	
Chronological age, months	38.6	42.1	49.4	3.5	7.3	10.8	
Sigma score							
Mean	.73	.88	.82	.15	— .06	.09	1.31
Standard deviation	.76	.71					
Intelligence quotient							
Mean	114.8	118.4	124.0	3.6	5.6	9.2	1.52
Standard deviation	16.1	14.4					
Percentile rank							
Mean	76.0	80.5	85.2	4.5	4.7	9.2	1.31
Standard deviation	23.3	21.0					
Binet intelligence quotient within one week							
Children	93	63	11				
Mean	122.6	124.3	129.8	1.7	5.5	7.2	.56
Standard deviation	15.6	20.8					
Binet intelligence quotient at entrance to preschool							
Children	93	62	10				
Mean	122.6	118.0	118.8	— 4.6	.8	— 3.8	1.76
Standard deviation	15.6	16.2					

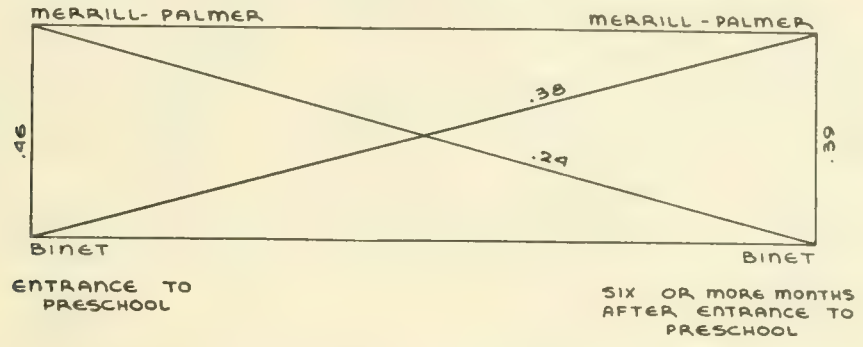
11 months attendance group did not deviate from this rule, however, is shown by their Binet IQ at entrance to preschool. In the 6 to 11 months of attendance they had gained 6.3 IQ points. This gain yielded a ratio of 1.91, or 97 chances in 100 of a true gain in Binet IQ during the 6 to 11 months of preschool attendance. The children who had attended sixteen or more months had gained 11.0 points in Binet IQ.

It will be seen that at entrance to preschool the different attendance groups were not equal in Binet IQ, the longer attendance groups being somewhat inferior on initial status. The 6 to 11 months group was 4.6 points inferior to the new entrants on initial status. This difference yielded a ratio of 1.76, or 96 chances in 100 of a true inferiority on initial status of the longer attendance group. This raises the question as to whether the same relative positions would have been found on Merrill-Palmer tests if the longer attendance group had received this test at entrance. There is no accurate way of knowing whether the longer attendance group was equal to the shorter attendance group at preschool entrance, slightly inferior to them or slightly superior to them in Merrill-Palmer scores.

The relationship between Merrill-Palmer IQ and Binet IQ for ninety-three children all of whom were new to preschool and tested 0 to 3 months after entrance was .46 ± .06. This relationship is not high as between two simultaneous (interval one week) measures of intelligence. The relationship is still lower when the children have been in attendance at preschool six or more months. The correlation for seventy-two such children having their first Merrill-Palmer test at this time and a simultaneous retest on the Binet was .39 ± .07. The correlation between first Merrill-Palmer IQ at this time and Binet IQ at the time of entrance to preschool (six or more months earlier) was .38 ± .07. The correlation between Merrill-Palmer IQ at entrance to preschool and Binet IQ after six months of preschool was still lower, being .24 ± .10 for forty-five children. It appears that the effect of preschool attendance on both Merrill-Palmer and Binet abilities is reflected in the lowered relationship between the two tests.

	Children	Correlation
Merrill-Palmer IQ at 0 to 3 months with simultaneous Binet IQ	93	.46 ± .05
Merrill-Palmer IQ at 6 or more months with simultaneous Binet IQ	72	.39 ± .07
Merrill-Palmer IQ at 6 or more months with Binet IQ at 0 to 3 months	72	.38 ± .07
Merrill-Palmer IQ at 0 to 3 months with Binet IQ at 6 or more months	45	.24 ± .10

The relationship may be expressed diagrammatically as follows:



Percentiles

There was an increase in mean percentile rank with increasing length of preschool attendance prior to the first test. The means were respectively 76.0, 80.5, and 85.2. The difference between the new entrants and the 6 to 11 months attendance group was 4.5 percentile ranks, which yielded a ratio of 1.31, or 90 chances in 100 of a true superiority for the longer attendance group. The difference between the 6 to 11 months attendance group and the sixteen or more months attendance group was 4.7 percentile ranks, making a total superiority of 9.2 percentile ranks of the longest attendance group over the shortest attendance group.

Age Differences

In Table 25 the attendance groups are divided according to age. The age groups varied considerably in ability at the time of entrance to preschool but the different methods of scoring are not in agreement as to which age groups were the superior ones. From the mean sigma scores it appears that the two- and four-year age groups were superior to the three- and five-year groups, while from the intelligence quotients and percentile ranks it appears that the four- and five-year-olds were superior to the two- and three-year-olds.

In comparing children of the same ages who had their first tests after different amounts of preschool attendance, the three-year-old children who were tested after six or more months of preschool attendance are found to be higher by each method of scoring than other three-year-old children who were tested on entrance to preschool. The five-year-old children show a slight tendency in the same direction of difference, while the reverse tendency is true of the four-year-old groups.

There appears to be no clear contribution to the problem from the analysis by age groups.

Table 25
Relation to Amount of Preschool Attendance Prior to First Test: Age Differences

Age, Months	Children	Mean Days Prior Attendance	Mean Chronological Age	Sigma Score		Intelligence Quotient		Percentile Rank	
				Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
0 to 3 Months Prior Attendance									
18 to 29	29	22.2	26.4	.91	.89	110.6	10.8	74.6	24.3
30 to 41	39	19.7	35.0	.49	.83	109.1	15.7	67.3	26.2
42 to 53	32	22.2	47.9	.92	.50	123.1	15.3	86.1	14.3
54 to 65	11	22.4	56.7	.55	.45	122.5	14.9	80.5	18.1
6 or More Months Prior Attendance									
30 to 41	35	114.6	35.8	.99	.81	116.8	13.6	80.0	20.8
42 to 53	29	133.1	46.9	.81	.62	120.3	15.0	82.4	19.6
54 to 65	11	194.2	56.6	.64	.37	124.3	13.1	81.8	18.2

Sex Differences

Table 26 shows that the mean score for girls was higher, although not significantly so, than for boys, both in entering groups and in groups who had attended preschool six or more months. The sixty-one girls tested at entrance to preschool were superior to the fifty boys in the following amounts: .11 sigma score, 4.6 points in IQ and 4.8 percentile ranks. These differences yielded ratios of .76, 1.6, and 1.1, or 77, 94, and 86 chances in 100 of true superiority of entering girls over entering boys.

The twenty-eight girls tested after six or more months of preschool attendance were superior to the forty-seven boys in the following amounts: .21 sigma score, 8.4 points in IQ and 8.0 percentile ranks.

Table 26
Relation to Amount of Preschool Attendance Prior to First Test: Sex Differences

Sex	Children	Mean Days Prior Attendance	Mean Chronological Age	Sigma Score		Intelligence Quotient		Percentile Rank	
				Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
0 to 3 Months Prior Attendance									
Boys	50	21.4	37.2	.67	.70	112.3	13.9	73.3	24.3
Girls	61	21.3	39.8	.78	.81	116.9	16.9	78.1	22.3
Difference		0.1	— 2.6	— .11		— 4.6		— 4.8	
Ratio				.76		1.6		1.1	
6 or More Months Prior Attendance									
Boys	47	118.2	42.2	.79	.65	116.1	13.3	78.2	20.4
Girls	28	158.9	44.8	1.00	.76	124.5	14.5	86.2	18.2
Difference		— 40.7	— 2.6	— .21		— 8.4		— 8.0	
Ratio				1.2		2.5		1.8	

The ratios were 1.2, 2.5, and 1.8, or 88, 99.4, and 96 chances in 100 of true superiority of the girls.

These findings suggest that girls may profit more than boys by preschool attendance. This conclusion should be considered tentative until similar results are obtained with larger groups.

Relation to Number of Days Preschool Attendance Prior to First Test

The children were also divided into groups according to the actual number of days they had attended preschool prior to their first Merrill-Palmer test. There were 108 children who had attended less than fifty days, 22 who had attended 50 to 99 days, 43 who had attended 100 to 149 days and 13 who had attended 150 or more days prior to their first Merrill-Palmer test (Table 27).

Sigma Scores

The children who had attended preschool 100 days or more prior to their first Merrill-Palmer test made slightly higher sigma scores than the children who had attended less than 50 days at the time of their first test. The difference was .15 and the ratio of the difference to the standard deviation of the difference was 1.15, indicating 87 chances in 100 of a true difference.

While the differences were not large and were not significant there was a tendency for children with more days of attendance prior to their test to receive higher sigma scores than the children who had been in attendance only a few days.

Intelligence Quotients

The mean Merrill-Palmer IQ of the 108 children who had attended preschool less than 50 days at the time of their first test was 114.6, while the mean IQ of the 56 children who had attended 100 days or more was 119.4, or 4.8 points higher. This difference yielded a ratio of 1.88, or 96 chances in 100 of a true difference.

The Binet intelligence quotients indicate that the group which attended the greater number of days had been slightly inferior to the shorter attendance group at the time they entered preschool. That is, before either group had attended as much as 50 days, the group which was to be in longer attendance at the time of the Merrill-Palmer test had been 3.3 Binet IQ points inferior. This difference yielded a ratio of 1.19, or 87 chances in 100 of a true inferiority. Whether these

Table 27
Relation to Number of Days Preschool Attendance Prior to First Test

Days Attendance	Children	Mean Chronological Age	Sigma Score		Intelligence Quotient		Percentile Rank		Binet Intelligence Quotient at Entrance		
			Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Children	Mean	Standard Deviation
1 to 49	108	38.4	.73	.77	114.6	15.6	75.8	23.5	90	122.4	15.8
50 to 99	22	43.8	.84		119.6		83.8		20	116.8	
100 to 149	43	41.2	.88		118.3		78.7		43	118.5	
150 or more	13	48.6	.85		122.9		85.2		12	121.5	
100 or more	56	43.0	.88	.78	119.4	15.5	80.2	22.4	55	119.1	16.3
Difference											
1 to 49 and	108	4.6	.15		4.8		4.4			3.3	
100 or more	56										
Ratio			1.15		1.88		1.17			1.19	

children were slightly inferior also in Merrill-Palmer ability at the time of entrance to preschool is not known.

Percentiles

The children who had attended 100 days or more at the time of their first Merrill-Palmer test were 4.4 percentile ranks superior to the children who had attended less than 50 days. This difference yielded a ratio of 1.17, or 87 chances in 100 of a true superiority.

Correlations

There was very little correlation between scores on the Merrill-Palmer test and number of days of preschool attendance prior to the test, except that the intelligence quotient for girls showed a relationship of .41 ± .06. The corresponding correlation for boys was .13 ± .07. The other two methods of scoring the Merrill-Palmer test showed little relationship for girls, the correlation for sigma score being .10 ± .07 and for percentile rank .18 ± .07. The correlations were:

	Boys (N = 95)	Girls (N = 89)
Sigma score	.06 ± .07	.10 ± .07
Intelligence quotient	.13 ± .07	.41 ± .06
Percentile rank	.09 ± .07	.18 ± .07

Summary

Several types of analyses were made to bring out the effects, if any, of preschool attendance upon Merrill-Palmer ability.

1. Changes during periods of preschool attendance compared with changes during periods of nonattendance:

a. Significant gains were made over the winter months when the children were enrolled in preschool. Over the summer vacation months, the changes were not significant. In sigma score there was a tendency to loss over the summer; in IQ and percentile some gains were made over the summer. In IQ the gain over the winter was approximately twice that over the summer. The difference in gains was not attributable to amount of previous experience with the test, initial status, age, or interval between tests.

b. The gains in IQ made over the winter months were inversely proportional to initial IQ level. The correlation of —.43 ± .07 contrasted with the positive correlation of .40 ± .10 on retests at an interval of only one week. These relationships indicate that the causes of change were not the same over the brief interval and the longer interval. The correlation over the summer months was —.38 ± .08. The cause of change on immediate retest appeared to be practice effects from recent experience with the test materials. Over the longer periods a new set of influences cancelled out the immediate effects of practice and produced a reversal of changes.

c. Changes in sigma score over winter and summer periods were inversely related to initial sigma score, in contrast to gains similar for all levels on retests at one week.

d. The changes served to reduce the correlations between tests. On retest at one week the correlations for the different methods of scoring (sigma score, IQ, and percentile) ranged from .89 to .96, at six months (summer) from .59 to .73,

and at six months (winter) from .48 to .60. The consistency over six months was greater in IQ than in sigma score or percentile rank.

e. The hypothesis that best fits the above facts is that a certain portion of the changes over the winter period was the resultant of environmental stimulation. Children of average ability found more stimulation in the preschool environment than in their summer home environments, responding by greater gains. Children of the highest levels of ability did not find the preschool environment sufficiently challenging to maintain their high level, although they maintained it to a greater degree than when not in preschool.

2. Changes according to number of days of attendance between tests:

a. There was no tendency for the changes during the winter to be related to the number of days of preschool attendance between tests. The correlations ranged from —.02 to —.16. A positive correlation would assume that the effects of preschool attendance, if any, work with the same rapidity for different children. Such was not the case in this group.

b. Changes in Binet IQ of the same children over the same period showed a slight positive relationship to number of days attendance, the correlation being $.25 \pm .08$.

3. Length of daily session:

Children attending preschool for half-day sessions gained more in Merrill-Palmer IQ than children attending all-day sessions. The reverse was true for Binet changes.

4. Interval between tests:

a. There were greater gains in IQ and in percentile with longer intervals between tests. No consistent trend of gain in sigma score was revealed.

b. Correlations between tests decreased with increasing length of interval. The correlations at one week, 6 months, 12 months, and 18 months were respectively: sigma score .92, .54, .50, and .47; IQ .90, .63, .52, and .50; percentile rank .89, .55, .26, and .49. The decrease should not be attributed to unreliability of the test alone, although uneven standardization of the test at different ages may have had some influence on the correlations.

5. Attendance prior to first test:

a. Children who had attended preschool six or more months prior to their first Merrill-Palmer test made higher scores than children new to preschool.

b. Girls made higher scores than boys. This was true of entering groups and of groups having attended preschool six or more months prior to their first test.

c. The relationship between Merrill-Palmer IQ and Binet IQ for children new to preschool was low ($.46 \pm .06$) as between two tests of intelligence. The relationship appeared to be lowered by preschool attendance, the correlation after six or more months of preschool being $.39 \pm .07$. The correlation between Merrill-Palmer IQ at this time (first test) and Binet IQ at entrance to preschool was $.38 \pm .07$, and between Merrill-Palmer IQ at entrance to preschool and Binet IQ after six or more months was $.24 \pm .10$. It thus appears that the effect of preschool attendance on both Merrill-Palmer and Binet abilities is reflected in the lowered relationships between the two tests.

6. Days attendance prior to first test:

a. There was a tendency for higher scores to be made by children who had attended more than 100 days prior to their first test than for children who had attended less than 50 days.

b. For boys there was no relationship between number of days attendance and scores on the test, the correlations being .06, .13, and .09. For girls there was no relationship when the scores used were sigma score and percentile, the correlations being .10 and .18, but a correlation of $.41 \pm .06$ was obtained for IQ, indicating higher IQ's for girls who had been in attendance more days prior to their first test.

CONCLUSIONS

From these various analyses a sufficient number of positive and significant differences was obtained to justify the conclusion that preschool attendance materially affected ability on the Merrill-Palmer test.

In general the IQ reflected these changes more clearly than sigma scores or percentiles.

The changes in Merrill-Palmer ability were less clean-cut than in Binet ability.

SOCIO-ECONOMIC STATUS

Occupation of Father

Relation of Scores on First Tests to Occupation of Father

Children enrolled in the preschool laboratories are drawn preponderantly from the professional classes. Approximately two-thirds of the fathers are so classified (5), while the other one-third are distributed in Classes II, III, and IV, according to the Goodenough (8) scheme of classification. The categories listed by Goodenough are:

Class I—Professional
Class II—Semiprofessional and managerial
Class III—Clerical, skilled trades, and retail business
Class IV—Semiskilled occupations, minor clerical positions, minor business, and farmers
Class V—Slightly skilled trades
Class VI—Day laborers

While the children enrolled in the preschools probably do not represent a random representative sample of preschool-age children whose fathers are classified in these occupations, it is nevertheless interesting to see what differences, if any, there are between children from different occupational classes when they enter preschool and also to see whether the differences, if found, persist after the children have had the benefits of preschool education in the laboratories. The preschools are open to children from all occupational levels, no attempt being made to govern admission on the basis of occupation.

On entrance to preschool, children from the professional classes have been found to have a significantly higher Binet IQ than children from Classes II to IV. The following figures are taken from Coffey and Wellman (5):

	Children	Mean IQ	Median IQ	Standard Deviation
Class I	278	119.4	118.4	16.6
Classes II to IV	139	112.0	111.7	14.9
Difference		7.4	6.7	
Ratio of difference to probable error of difference		6.85	4.96	

However, the changes in Binet IQ from fall to spring during the first year of preschool were very similar, the mean gain for Class I children being 7.1 points, with a standard deviation of 11.6, and for children in Classes II to IV 6.1 points, with a standard deviation of 11.3. The ratio of the difference in gains (1.0 point) to the probable error of the difference was 1.25. Gains for each of the four classes were from 5 to 7 points.

Table 28
Relation of Scores on Merrill-Palmer First Test to Occupation of Father

Occupation	Children	Sigma Score		Intelligence Quotient		Percentile Rank		Binet Intelligence Quotient		
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Children	Mean	Standard Deviation
All First Tests										
Class I	156	.79	.76	116.7	15.9	77.8	22.6	134	125.8	19.0
Classes II to IV	52	.79	.71	117.6	15.6	79.3	21.1	47	119.2	13.2
Difference		.00		— .9		— 1.5			6.6	
Ratio of difference to standard deviation of difference									2.6	
0 to 3 Months Attendance Prior to First Test										
Class I	81	.68	.80	113.7	16.3	73.6	24.9	65	122.5	16.4
Classes II to IV	30	.88	.64	117.9	14.5	82.0	16.9	28	122.6	13.4
Difference		— .20		— 4.2		— 8.4			— .1	
Ratio of difference to standard deviation of difference		1.3		1.3		2.0			.03	
6 or More Months Attendance Prior to First Test										
Class I	56	.92	.66	120.0	13.4	82.9	17.3	52	129.9	20.1
Classes II to IV	21	.69	.78	118.0	17.3	76.7	25.3	18	113.6	10.9
Difference		.23		2.0		6.2			16.3	
Ratio of difference to standard deviation of difference		1.2		.47		1.0			4.3	

Neither Class I nor Classes II to IV gained over the summer months, the change for Class I being —0.62 point and for Classes II to IV 0.30 point.

Table 28 compares the scores on the first Merrill-Palmer test according to occupational status. The second section of the table, which presents separately the results for new entrants to preschool, can be compared directly with the data just quoted on Binet results. The third section of the table presents results for children who had their first test after they had attended preschool at least six months. This group, in comparison with the new entrants, was intended to give an indication of how the preschools have affected children coming from different occupational classifications.

Sigma Scores.—There was no difference in sigma score mean between the two occupational groupings. The fifty-two preschool children whose fathers were in occupations II to IV made the same mean sigma score on their first test as that made by the 156 children whose fathers were in the professions.

New entrants from Classes II to IV were slightly superior in mean sigma score to new entrants from Class I. The difference was —.20 and the ratio 1.3 or 90 chances in 100 of a true inferiority of children from the professional classes. Children from the professional classes who had attended preschool for several months prior to their test were superior to children from the lower classes, however. The difference was .23 sigma score and the ratio 1.2 or 88 chances in 100 of a true superiority. The interpretation of these findings in terms of the effects of preschool attendance is not clear, since the status on entrance to preschool of the group tested after several months of attendance is not known. If the status of this group at entrance is assumed as equal to the other new entrants (probably a questionable assumption) then it would appear that preschool attendance increased the ability of children from the professional classes (The difference between new entrants and those with prior attendance is .24, the ratio 2.0, and chances in 100 are 98.), while at the same time it tended to decrease the ability of children from the lower classes (The difference between new entrants and those with prior attendance is —.19, the ratio .9, and chances in 100 are 82.). Probably a safer conclusion to draw is that stable results on the Merrill-Palmer test in relation to occupational level should not be expected from groups with such small numbers of cases as are represented here, particularly in the lower classes when

subdivided according to attendance. Sampling errors may overshadow other factors.

Intelligence Quotients.—The fifty-two children whose fathers were in occupations II to IV had a mean IQ .9 of a point higher than the 156 children whose fathers were in the professions. This lack of difference between the occupational classes is interesting in comparison with the simultaneous Binet tests, which show a superiority of 6.6 IQ points for the professional classes. The difference in Binet IQ for these children is almost identical with the difference for the larger groups studied by Coffey and Wellman (5), indicating that the relatively small number of cases in the lower occupational groups does not constitute an explanation for the lack of difference in Merrill-Palmer IQ. The ratio of the difference in Binet IQ to the standard deviation of the difference was 2.6, or 99.5 chances in 100 of a true superiority in Binet IQ of the children from the professional classes.

New entrants to preschool from the professional classes were 4.2 points inferior in Merrill-Palmer IQ to new entrants from the lower classes. Children from the professional classes who had attended preschool several months prior to their first test were 2.0 points superior to children from the lower classes with similar attendance, however. The explanation for these conflicting tendencies appears to lie in the small number of cases compared when the groups are subdivided according to amount of attendance.

In Binet IQ new entrants from the two classes were equal, the difference being only .1 point; while children from the professional classes with six or more months of preschool attendance were 16.3 points superior to children of equal attendance from classes II to IV. It will be noted that these new entrants from the lower classes represented a superior selection in Binet ability, since they were 10 points higher in IQ than the larger group studied by Coffey and Wellman (5).

Children from the professional classes who had attended preschool six or more months were 6.3 Merrill-Palmer IQ points superior to new entrants to preschool from the professional classes. The ratio was 2.5, or 99.4 chances in 100 of a true superiority. There was no difference in Classes II to IV between new entrants and children who had had several months of preschool attendance. In Binet IQ, children from the professional classes who had prior attendance were also superior to new entrants from the professional classes. The difference was 7.4 points, the ratio 2.1, and the chances in 100 were 98. However, chil-

dren from classes II to IV with six or more months of attendance were 9.0 points inferior in Binet IQ to new entrants from these classes. The ratio was 2.5, or 99.4 chances in 100 of a true inferiority. These results do not check with the data on larger numbers of cases in the Coffey and Wellman (5) study.

In view of the conflicting tendencies, it seems unwise to draw conclusions on the small number of cases represented when the groups are subdivided according to amount of attendance.

Percentiles.—No difference between the occupational groups was revealed by percentile ranks. The difference in the means was 1.5 ranks in favor of the lower occupational groups. When the groups were subdivided according to amount of preschool attendance prior to first test, the same reversal of tendencies was evident as for sigma score and intelligence quotient. These conflicting tendencies appear to be sampling errors.

Sex Differences in Scores Related to Occupation of Father

In Binet IQ in the Coffey and Wellman (5) study, girls were higher than boys, although the difference was not completely significant. Girls of Class I occupation were 4.3 points higher in mean Binet IQ than boys of Class I. The ratio of the difference to the probable error of the difference was 3.9, or 99.6 chances in 100 of a true difference. Girls of Classes II to IV also had a mean 4.3 points higher than the boys of these classes. The ratio was 2.9, or 97.5 chances in 100 of a true difference.

Sex	Children	Mean Binet IQ	Standard Deviation
		Class I	
Girls	193	122.5	15.4
Boys	219	118.2	17.3
		Classes II to IV	
Girls	92	115.8	15.0
Boys	88	111.5	14.9

Boys of Class I were significantly higher than boys of the lower classes, the difference being 6.7 points and the ratio 5.1 Girls of Class I were also significantly higher than girls of the lower classes, the superiority being 6.7 points, and the ratio 5.2.

Although there were these differences between the sexes in mean IQ when the children entered preschool, the gains were similar for the sexes. The girls of Class I gained 6.0 points from fall to spring, the boys of this class 6.2 points. The girls of Classes II to IV gained 5.7 points and the boys of these classes 4.9 points.

Neither sex changed significantly over the summer when not in

preschool. The girls of Class I lost 1.6 points, the boys gained 0.1 point. The girls of Classes II to IV lost 0.7 points and the boys gained 1.4 points.

Table 29 presents the scores on Merrill-Palmer first tests for the two sexes divided according to occupation of father.

Sigma Scores.—The seventy-two girls whose fathers were in the professions had a mean sigma score .18 higher than the eighty-four boys whose fathers were in the professions. This difference yielded a ratio of 1.5, or 93 chances in 100 of a true superiority of the girls. The twenty-seven girls whose fathers were in occupational classes II to IV, however, were .14 sigma score inferior to the twenty-five boys from these classes. The ratio was .70, indicating 76 chances in 100 of a true superiority of the boys.

The girls from the professional classes were .16 sigma score superior to the girls from the lower occupational groups. The boys from the professional classes were, however, .16 sigma score inferior to the boys from the lower occupational groups. Neither difference was significant, the ratio in each case being 1.0, or 84 chances in 100 of a true difference.

Intelligence Quotients.—The girls from the professional classes were 5.7 points superior in Merrill-Palmer IQ to the boys from the professional classes. This difference yielded a ratio of 2.3, or 98.9 chances in 100 of a true superiority of the girls. Their superiority in Binet IQ was 4.6 points, which corresponded very closely to the 4.3 points difference for larger numbers of children in the Coffey and Wellman (5) study.

Girls from occupational classes II to IV were slightly inferior to boys from occupational classes II to IV, however, in both Merrill-Palmer and Binet IQ. The difference was only 1.0 point in Merrill-Palmer IQ and can be considered as no essential difference. The difference in Binet IQ was 2.5 points. The direction of difference here does not correspond to the direction of difference for the larger groups studied by Coffey and Wellman (5). It seems probable, therefore, that the number of cases is too small in these occupations to make comparisons between the sexes reliable and representative.

The girls from the professional classes were slightly superior to the girls from the lower classes in Merrill-Palmer and considerably superior in Binet IQ, the differences being 2.7 and 10.5 IQ points respectively. The ratio of the difference was .7 in Merrill-Palmer and 2.8

Table 29
Sex Differences in Scores Related to Occupation of Father

Sex Group	Children	Sigma Score		Intelligence Quotient		Percentile Rank		Binet Intelligence Quotient		
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Children	Mean	Standard Deviation
Occupational Class I										
Girls	72	.88	.82	119.8	17.4	81.1	22.0	62	128.4	18.7
Boys	84	.70	.68	114.1	14.1	75.3	23.1	72	123.8	19.2
Difference		.18		5.7		5.8			4.6	
Ratio of difference to standard deviation of difference		1.5		2.3		1.6			1.4	
Occupational Classes II to IV										
Girls	27	.72	.71	117.1	15.8	78.6	21.1	23	117.9	14.0
Boys	25	.86	.70	118.1	15.4	80.1	21.6	24	120.4	12.2
Difference		— .14		— 1.0		— 1.5			— 2.5	
Ratio of difference to standard deviation of difference		.70							0.7	

in Binet IQ. The latter ratio indicated 99.7 chances in 100 of a true superiority in Binet IQ of girls in the professional classes over girls in the lower classes.

The boys from the professional classes were, however, 4.0 points inferior in Merrill-Palmer IQ to the boys from the lower classes. This difference yielded a ratio of 1.2, or 88 chances in 100 of a true inferiority. In Binet IQ, the boys from the professional classes were 3.4 points superior to the boys from the lower classes. The ratio was 1.0, or 84 chances in 100 of a true superiority.

Percentiles.—The girls from the professional classes were 5.8 percentile ranks superior to the boys from the professional classes. The ratio was 1.6, indicating 94 chances in 100 of a true superiority. The girls from the lower occupational classes were 1.5 percentile ranks inferior to the boys from the lower occupational classes. This difference may be considered insignificant.

The girls from the professional classes were 2.5 percentile ranks superior to the girls from the lower classes. The ratio was .6, or 73 chances in 100 of a true superiority.

The boys from the professional classes were 4.8 percentile ranks inferior to the boys from the lower classes. The ratio was 1.0 or 84 chances in 100 of a true inferiority.

Education of Father

Relation of Scores on First Tests to Education of Father

Table 30 presents scores on first tests when the children were classified according to the number of years of formal schooling of the father. Tables 31 and 32 present similar material according to schooling of the mother and according to midparent education, the latter being the average of father's and mother's education for each child. The classification 8 to 15 years represent grammar school education only through graduation from high school and three years of college. The classification 16 to 20 years represents graduation from college and further time spent in higher education.

In the study by Coffey and Wellman (5) on the Binet test it was possible to subdivide the first classification into 8 to 11 years and 12 to 15 years. The upper limit of education in the first group thus represented high school education with failure to complete the high school course. On the Binet test children whose fathers had had more

Table 30

Relation of Scores on Merrill-Palmer First Tests to Education of Father

Education, Years	Children	Sigma Score		Intelligence Quotient		Percentile Rank		Binet Intelligence Quotient		
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Children	Mean	Standard Deviation
All First Tests										
8 to 15	40	.68	.88	116.0	19.3	73.8	26.6	37	119.0	15.2
16 to 20	156	.80	.70	116.9	14.6	79.0	20.9	121	124.7	18.9
Difference		.12		.9		5.2			5.7	
Ratio of difference to standard deviation of difference		.80		.28		1.2			1.9	
0 to 3 Months Attendance Prior to First Test										
8 to 15	21	.55	.82	114.0	20.0	71.0	26.4	19	121.3	15.6
16 to 20	81	.75	.74	114.7	14.7	76.6	23.0	66	122.4	16.2
Difference		.20		.7		5.6			1.1	
Ratio of difference to standard deviation of difference		1.00		.15		.89			.27	
6 or More Months Attendance Prior to First Test										
8 to 15	18	.72	.86	117.2	18.5	75.6	26.9	18	116.7	13.3
16 to 20	56	.88	.62	119.6	12.7	82.6	17.3	55	127.4	21.3
Difference		.16		2.4		7.0			10.7	
Ratio of difference to standard deviation of difference		.73		.53		1.04			2.5	

Table 31

Relation of Scores on Merrill-Palmer First Tests to Education of Mother

Education, Years	Children	Sigma Score		Intelligence Quotient		Percentile Rank		Binet Intelligence Quotient		
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Children	Mean	Standard Deviation
All First Tests										
8 to 15	84	.78	.78	117.6	16.8	77.8	23.8	68	121.4	15.7
16 to 20	111	.78	.72	116.4	15.3	78.1	21.3	89	125.1	19.6
Difference		.00		— 1.2		0.3			3.7	
Ratio of difference to standard deviation of difference				.51		.09			1.3	
0 to 3 Months Attendance Prior to First Test										
8 to 15	42	.69	.80	115.0	16.7	74.6	24.9	33	120.5	14.5
16 to 20	60	.75	.74	114.9	16.2	75.5	23.1	52	123.0	17.1
Difference		.06		— 0.1		0.9			2.5	
Ratio of difference to standard deviation of difference		.38		.03		.19			.71	
6 or More Months Attendance Prior to First Test										
8 to 15	35	.80	.72	118.8	15.4	80.2	22.0	35	122.2	16.8
16 to 20	38	.89	.66	119.3	13.5	81.5	18.7	37	128.0	22.3
Difference		.09		0.5		1.3			5.8	
Ratio of difference to standard deviation of difference		.56		.15		.27			1.3	

Table 32
Relation of Scores on Merrill-Palmer First Tests to Midparent Education

Education, Years	Children	Sigma Score		Intelligence Quotient		Percentile Rank		Binet Intelligence Quotient		
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Children	Mean	Standard Deviation
All First Tests										
8 to 15	62	.71	.82	115.9	17.9	75.0	24.7	53	118.8	15.8
16 to 20	132	.80	.72	117.0	14.7	79.2	21.1	103	125.9	18.9
Difference		.09		1.1		4.2			7.1	
Ratio of difference to standard deviation of difference		.75		.42		1.15			2.5	
0 to 3 Months Attendance Prior to First Test										
8 to 15	35	.63	.78	114.2	17.6	73.0	25.3	29	119.9	14.9
16 to 20	66	.76	.76	114.6	15.1	76.7	23.1	55	123.1	16.9
Difference		.13		0.4		3.7			3.2	
Ratio of difference to standard deviation of difference		.81		.11		.72			.89	
6 or More Months Attendance Prior to First Test										
8 to 15	24	.73	.83	116.2	17.1	76.1	24.5	24	117.4	16.6
16 to 20	49	.91	.60	120.5	12.8	83.2	17.5	48	129.1	20.4
Difference		.18		4.3		7.1			11.7	
Ratio of difference to standard deviation of difference		.95		1.06		1.27			2.6	

years of schooling had higher intelligence quotients. The mean IQ's were as follows:

Father's Education, Years	Children	Mean IQ	Standard Deviation
8 to 11	16	106.4	12.5
12 to 15	114	116.9	15.0
16 to 20	347	119.6	16.9

The difference between the first two classifications (8 to 11 years and 12 to 15 years) was statistically significant, the difference being 10.5 points in favor of the group whose fathers were better educated, and the ratio of the difference to the probable error of the difference being 4.6. The children whose fathers had 16 to 20 years of education were 2.7 points in IQ superior to the children whose fathers had 12 to 15 years of education, but the difference was not statistically significant.

Sigma Scores.—The children whose fathers had 16 to 20 years of schooling were only .12 sigma score on Merrill-Palmer superior to the children whose fathers had less formal schooling. New entrants to preschool whose fathers had the greater amount of schooling made sigma scores .20 superior to children whose fathers were less educated. The superiority was also evident after six or more months of preschool attendance. None of the differences were completely significant, however, the ratios being .80, 1.00, and .73 respectively, with 79, 84, and 76 chances in 100 of true superiority of the children whose fathers were better educated.

Expressed in correlations (Table 33) no relationship was found between sigma score and father's education either for boys or girls, for new entrants, or for children with prior attendance. The correlations ranged from .04 ± .10 to .17 ± .09.

It should be kept in mind that most of the fathers even in the lower group represented a somewhat superior selection educationally since all but six of the forty had graduated from high school.

Intelligence Quotients.—There was no difference in Merrill-Palmer IQ between the children whose fathers were better educated and the children whose fathers had less education. The differences were .9, .7, and 2.4 points respectively and the ratios .28, .15, and .53. The correlations (Table 33) ranged from −.11 ± .09 to .14 ± .09. In Binet IQ both groups were superior to the larger groups studied by Coffey and Wellman (5) but the children whose fathers were better educated were 5.7 points higher than the children whose fathers were less well educated. The ratio was 1.9, indicating 97 chances in 100 of a true superiority in Binet IQ of children whose fathers were better

Table 33
Correlations Between Children's Scores and Education of Parents

Education	Children	Sigma Score	Intelligence Quotient	Percentile Rank	Binet	
					Children	Intelligence Quotient
Scores at 0 to 3 Months of Preschool Attendance						
Girls and Boys						
Father	102	.04 ± .10	—.04 ± .10	.06 ± .07	85	.07 ± .08
Mother	102	.12 ± .07	.03 ± .07	.08 ± .07	85	.14 ± .08
Midparent	101	.10 ± .07	—.02 ± .07	.06 ± .07	84	.10 ± .08
Girls						
Father	56	.07 ± .09	.04 ± .10	.06 ± .09	45	.16 ± .10
Mother	57	.06 ± .09	.08 ± .09	—.02 ± .10	46	.07 ± .09
Midparent	56	.10 ± .09	.06 ± .09	—.02 ± .10	45	.18 ± .10
Boys						
Father	46	.11 ± .09	—.11 ± .09	.07 ± .09	40	—.03 ± .11
Mother	45	.14 ± .10	—.01 ± .11	.21 ± .10	39	.21 ± .10
Midparent	45	.14 ± .10	—.09 ± .09	.12 ± .10	39	.05 ± .11
Scores After 6 or More Months of Preschool Attendance						
Girls and Boys						
Father	74	.10 ± .08	.05 ± .08	.12 — .08	73	.18 ± .08
Mother	73	.20 ± .08	.08 ± .08	.20 — .08	72	.30 ± .07
Midparent	73	.14 ± .08	.05 ± .08	.16 ± .08	72	.27 ± .07
Girls						
Father	26	.07 ± .12	.02 ± .12	.21 ± .12	26	.27 ± .14
Mother	25	.28 ± .14	.21 ± .14	.44 ± .13	25	.42 ± .13
Midparent	25	.12 ± .15	.05 ± .15	.28 ± .14	25	.35 ± .13
Boys						
Father	48	.17 ± .09	.14 ± .09	.13 ± .09	47	.20 ± .09
Mother	48	.20 ± .09	.09 ± .09	.11 ± .09	47	.28 ± .09
Midparent	48	.25 ± .09	.16 ± .09	.01 ± .09	47	.26 ± .09

educated. New entrants from the lesser educated groups were exceptionally high when compared with the same levels in the Coffey and Wellman (5) group. Correlations between education of fathers and Binet IQ's of children were negligible, ranging from —.03 to .27 ± .14.

Percentiles.—Comparisons on percentile ranks indicated a trend toward superiority of the children whose fathers were better educated. The children whose fathers had 16 to 20 years of education were 5.2 percentile ranks superior to the children whose fathers had 8 to 15 years of schooling. The ratio was 1.2 or 88 chances in 100 of a true superiority. New entrants to preschool with better educated fathers tended to be superior to new entrants with fathers of less schooling, the difference being 5.6 percentile ranks, the ratio .89, and the chances of true superiority 81 in 100. The trend toward superiority of the better educated persisted after preschool attendance, the difference being 7.0 percentile ranks, the ratio 1.04, and the chances 85 in 100 of a true superiority. Correlations ranged from .06 to .21 ± .12.

Education of Mother

Relation of Scores on First Tests to Education of Mother

In Binet IQ, as reported by Coffey and Wellman (5) the children whose mothers had 12 to 15 years of education were significantly higher than those whose mothers had had only 8 to 11 years of schooling. The difference was 14.8 points and the ratio of the difference to the probable error of the difference was 5.5. The children whose mothers had had 16 to 20 years of schooling were not significantly higher than those whose mothers had 12 to 15 years of education.

Mother's Education, Years	Mean IQ	Children	Standard Deviation
8 to 11	10	103.0	12.1
12 to 15	214	117.8	16.8
16 to 20	253	119.7	16.3

The gains made in Binet IQ from fall to spring in the Coffey and Wellman (5) study were similar for the three educational groups, however, there being no significant difference. The gains were, respectively, 5.4, 7.1, and 5.9 points in IQ.

Table 31 presents the scores on Merrill-Palmer tests in relation to number of years of mother's education.

Sigma Scores. There was no difference in sigma scores between children whose mothers had more education and children whose mothers had less education. The means were identical. For new entrants the difference was .06 and for children with prior preschool attendance .09. The latter yielded a ratio of .56 or only 71 chances in 100 of a true difference. Correlations (Table 33) showed no relationship, ranging from .06 to .28 ± .14.

Intelligence Quotients.—No difference in intelligence quotient was found between the groups classified according to education of mother. The differences were respectively —1.2, —0.1, and 0.5 points. Correlations ranged from —.01 to .21 ± .14. In Binet IQ the children were higher than comparable groups in the Coffey and Wellman (5) study. The children whose mothers were better educated showed some superiority over those whose mothers were less educated, but the difference was not statistically significant. The superiority was 3.7 points, the ratio 1.3, and the chances in 100 of a true superiority were 90. Correlations between mother's education and Binet IQ ranged from .07 to .42 ± .13.

Percentile Ranks.—There was no tendency towards higher percentile ranks earned by children whose mothers were better educated. The

mean differences were 0.3, 0.9, and 1.3. Correlations ranged from —.02 to .44 ± .13.

Midparent Education

Relation of Scores on First Tests to Midparent Education

There was in some cases considerable difference between the education of father and mother. The child's score was therefore related to the average of father's and mother's education, termed midparent education. When this was done in the Coffey and Wellman (5) study, the Binet IQ's were significantly different for the three classifications. Midparent education revealed differences between the two higher educational classifications that were not shown when each parent was considered separately. The following differences were found:

Midparent Education, Years	Children	Mean IQ	Standard Deviation
8 to 11	20	105.4	13.0
12 to 15	181	115.8	15.3
16 to 20	263	121.2	16.6

The children whose midparent education was 12 to 15 years had IQ's 10.4 points higher than those whose midparent education was only 8 to 11 years. The ratio of the difference to the probable error of the difference was 5.2. Children whose midparent education was 16 to 20 years were significantly higher than those whose midparent education was 12 to 15 years. The difference was 5.4 points and the difference divided by the probable error of the difference was 4.9. The children whose midparent education was 16 to 20 years entered preschool with a Binet IQ 15.8 points higher than those whose midparent education was only 8 to 11 years.

Although there were distinct differences in the mean Binet IQ with different levels of midparent education, the changes in IQ in the Coffey and Wellman (5) study were not significantly different for the three classifications. During a preschool year (from fall to spring) the changes were respectively 6.9, 6.8, and 5.8 points gain.

Table 32 presents the Merrill-Palmer scores in relation to midparent education.

Sigma Scores.—There was little difference in the mean sigma scores earned by the children with higher midparent education and the children with lower midparent education. The differences were .09, .13, and .18 and the ratios were .75, .81, and .95. Correlations (Table 33) between midparent education and child's sigma score ranged from .10 ± .07 to .25 ± .09.

Intelligence Quotients. There was no difference in intelligence quo-

tients for the group as a whole or for the new entrants. Children who had had six or more months of preschool attendance and whose midparent education was 16 to 20 years showed a 4.3 point superiority over other children of similar preschool attendance whose midparent education was 8 to 15 years. The ratio was 1.06 and the chances 85 in 100 of a true superiority. This would seem to suggest the hypothesis that the combination of preschool attendance and high midparent education may be an effective one in lifting the Merrill-Palmer IQ level of the children. The number of cases in the subdivided groups is small, however, and the suggestion should be considered as tentative. Correlations (Table 33) between midparent education and child's IQ ranged from —.09 to .16.

In Binet IQ children with higher midparent education were 7.1 points superior to children with lower midparent education. The ratio was 2.5 or 99.4 chances in 100 of a true superiority. Both groups were high in comparison with the larger group in the Coffey and Wellman (5) study. The difference between new entrants was 3.2 points and between children with prior attendance 11.7 points. New entrants with lower midparent education were unusually high in comparison with the Coffey and Wellman group. Correlations between midparent education and child's Binet IQ ranged from .05 to .35 ± .13. They indicate that while for large groups differences in means may be found, individual variations are so great that correlations are negligible.

Percentiles.—Children whose midparent education was 16 to 20 years were 4.2 percentile ranks superior to children whose midparent education was 8 to 15 years. The ratio was 1.15, or 87 chances in 100 of a true superiority. New entrants with the greater amount of midparent education were 3.7 percentile ranks superior to new entrants with the lesser amount of midparent education. The ratio was .72, or 76 chances in 100 of a true superiority. Children with six or more months of preschool attendance prior to their first test whose midparent education was 16 to 20 years were 7.1 percentile ranks superior to other children with equal prior preschool attendance whose midparent education was 8 to 15 years. The ratio was 1.27, or 89 chances in 100 of a true superiority. Correlations ranged from —.02 to .28 ± .14.

Summary

Occupation of Father

Fathers of children attending the preschool laboratories are preponderantly in the professional classes. However, about one-fourth of

the children examined on the Merrill-Palmer test had fathers in Classes II to IV according to the Goodenough scheme of classification. Comparison of the scores made on their first Merrill-Palmer test did not reveal differences between the children whose fathers were in Class I and the children whose fathers were in Classes II to IV.

This lack of difference between the occupational classes is particularly interesting in view of the fact that simultaneous (interval one week) Binet tests on the same children revealed a superiority of 6.6 points in IQ in favor of the professional classes. The ratio on Binet IQ was 2.6 or 99.5 chances in 100 of a true superiority. Further, the amount of superiority in Binet IQ for this group corresponded very closely to the results in a previous study for a considerably larger number of preschool children, for whom the difference in Binet IQ was 6.7 points and was statistically significant.

Education of Parents

Children were classified according to number of years of formal schooling of their fathers and their mothers and according to midparent education. No outstanding differences in Merrill-Palmer ability were revealed between children whose parents were better educated and children whose parents were less well educated. There was a slight tendency for children whose parents had 16 to 20 years of schooling to score higher than children whose parents had only 8 to 15 years of schooling, although the differences were not significant.

The lack of differences is again somewhat in contrast with Binet results on the same children. While the differences in Binet IQ were not entirely significant, they more nearly approached statistical significance. The superiority of the children of better educated parents (16 to 20 years of formal schooling) over children of less educated parents (8 to 15 years of schooling) is expressed by the following ratios: fathers 1.9, mothers 1.3, and midparent 2.5. The chances in 100 of true differences were 97, 90, and 99.4 respectively. In the larger group previously studied by Coffey and Wellman (5) significant differences in Binet IQ were found for midparent education of three groupings: 8 to 11 years, 12 to 15 years, and 16 to 20 years.

For the children of this study there was no appreciable correlation between education of parents and scores, either for boys or girls, for new entrants to preschool, or for children with six or more months of prior preschool attendance. The correlations ranged from —.11 ± .09 to .44 ± .13. Binet correlations were insignificant also, ranging from —.03 to .42 ± .13.

RELATION TO BINET

Correspondence Between Merrill-Palmer and Binet on Tests One Week Apart

Three hundred and forty children were tested on both Merrill-Palmer and Binet scales at an interval not greater than one week, in rotated order, some receiving the Merrill-Palmer test first and some the Binet first. Two- and three-year-old children were tested on the Kuhlmann revision and four- and five-year-old children were tested on the Stanford revision.

Correspondence Between Intelligence Quotients

The correspondence between Merrill-Palmer and Binet intelligence quotients is expressed in terms of correlations by age in Table 34 and in terms of mean differences by age and mean differences by IQ levels and mental age levels in Tables 34, 35, and 36.

When Grouped According to Age.—The correlations between Merrill-Palmer and simultaneous Binets were not high as between two different tests of intelligence. They were high enough (.45 to .63) to indicate some communality between the two measures but not high enough for predictive purposes. There was no tendency for the size of the correlation to increase or decrease proportionately to age, the highest correlation of the series being for three-year-olds and the lowest for four-year-olds.

The mean Merrill-Palmer IQ rose noticeably with age, but the Binet means were fairly close together at the different ages and tended to be higher than the corresponding Merrill-Palmers. Consequently the difference between the two tests was greatest at the younger ages. The discrepancy between means did not affect the correlations. For the group as a whole (340 cases) the Merrill-Palmer mean was 5.3 points lower than the Binet mean.

The above correlations were computed on simultaneous tests regardless of whether they were first or later tests and regardless of the amount of preschool attendance. The other two correlations in Table 34 taken from *Effects of Preschool Attendance* (p. 70) show that groups selected on different bases (in these two cases all were first tests on the Merrill-Palmer, but one group was limited to new entrants and the other limited to children with prior preschool attendance) may

Table 34

Correspondence Between Merrill-Palmer and Binet Intelligence Quotients on Simultaneous Tests (One Week Interval) When Grouped According to Age

Age, Months	Children	Merrill-Palmer Intelligence Quotient		Binet Intelligence Quotient			Difference in Intelligence Quotient	r	P.E.
		Mean	Standard Deviation	Test	Mean	Standard Deviation			
Age Groups									
18 to 29	32	109.5	12.1	Kuhlmann	121.6	15.5	—12.1	.50 ±	.09
30 to 41	126	114.6	16.4	Kuhlmann	123.3	21.7	— 8.7	.63 ±	.04
42 to 53	125	125.3	17.8	Stanford	128.8	17.0	— 3.5	.45 ±	.05
54 to 65	57	124.6	15.7	Stanford	122.3	15.7	2.3	.54 ±	.06
18 to 65	340	119.7			125.0		— 5.3	.52 ±	.03
0 to 3 Months Attendance Prior to First Test									
18 to 65	93			Kuhlmann and Stanford				.46 ±	.06
6 or More Months Attendance Prior to First Test									
18 to 65	72			Kuhlmann and Stanford				.39 ±	.07

Table 35

Correspondence Between Merrill-Palmer and Binet Intelligence Quotients on Tests One Week Apart When Grouped According to Merrill-Palmer IQ or Binet IQ Level

IQ Level	Children	Merrill-Palmer Intelligence Quotient	Binet Intelligence Quotient	Difference
Merrill-Palmer				
50 to 79	3	71.0	79.7	+ 8.7
80 to 89	9	85.0	92.7	+ 7.7
90 to 99	27	94.2	110.6	+ 16.4
100 to 109	66	104.6	118.0	+ 13.4
110 to 119	74	114.7	125.2	+ 10.5
120 to 129	57	124.9	129.6	+ 4.7
130 to 139	57	134.2	132.2	— 2.0
140 to 149	28	144.0	135.9	— 8.1
150 to 159	11	153.1	138.2	— 14.9
160 to 179	2	166.0	154.0	— 12.0
Binet				
50 to 79	6	85.8	74.2	— 11.6
80 to 89	3	87.3	87.0	— 0.3
90 to 99	14	100.9	96.0	— 4.9
100 to 109	49	112.1	104.7	— 7.4
110 to 119	48	114.8	114.3	— 0.5
120 to 129	81	119.8	125.3	+ 5.5
130 to 139	61	125.3	134.0	+ 8.7
140 to 149	39	129.6	144.2	+ 14.6
150 to 159	24	135.1	153.9	+ 18.8
160 to 179	9	127.5	167.2	+ 39.7

Table 36

Correspondence Between Merrill-Palmer and Binet Mental Ages on Tests One Week Apart

Mental Age, Months	Children	Merrill-Palmer Mental Age	Binet Mental Age	Difference
Grouped According to Merrill-Palmer Mental Age				
18 to 29	21	25.9	29.6	+ 3.7
30 to 41	80	35.4	39.3	+ 3.9
42 to 53	74	46.7	51.4	+ 4.7
54 to 65	81	57.9	60.4	+ 2.5
66 to 77	52	72.6	69.2	— 3.4
78 to 89	19	78.0	73.1	— 4.9
Grouped According to Binet Mental Age				
18 to 29	14	26.5	26.5	0.0
30 to 41	66	33.8	35.6	+ 1.8
42 to 53	83	46.2	48.2	+ 2.0
54 to 65	87	57.8	58.8	+ 1.0
66 to 77	62	68.2	70.9	+ 2.7
78 to 89	15	72.3	80.9	+ 8.6

yield different figures, the correlations being respectively .46 ± .06 and .39 ± .07. The latter figure is below the correlation obtained for the age groups and for the total group.

When Grouped According to Merrill-Palmer IQ Level.—The children were grouped according to Merrill-Palmer intelligence quotient, by ten-point divisions, and the mean Binet intelligence quotient was computed for each division. The results are presented in Table 35 and Figures

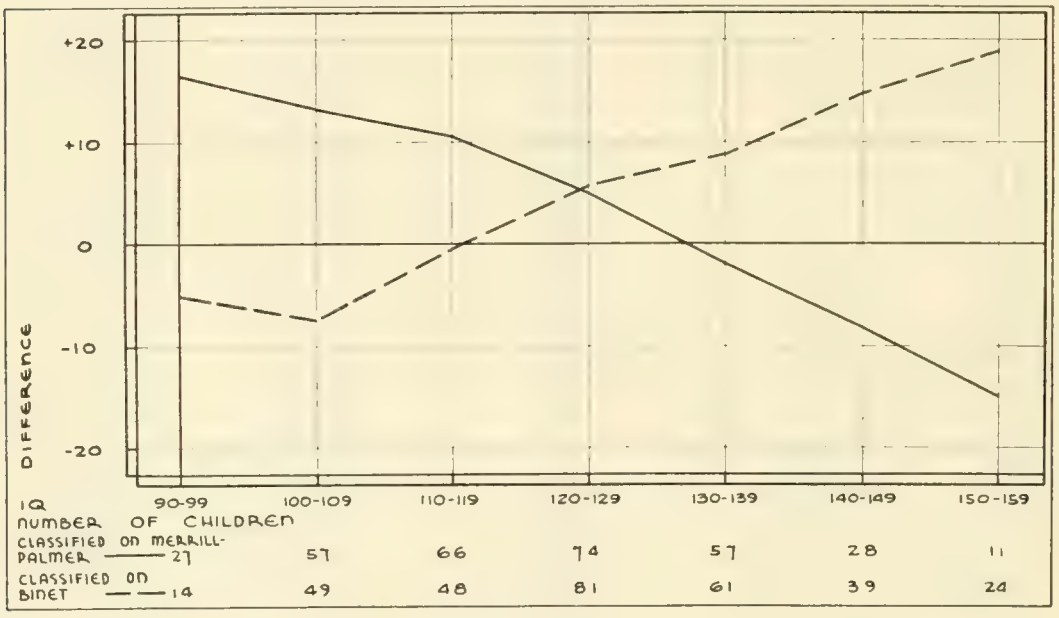


Figure 9. Divergence of Binet IQ From Merrill-Palmer IQ

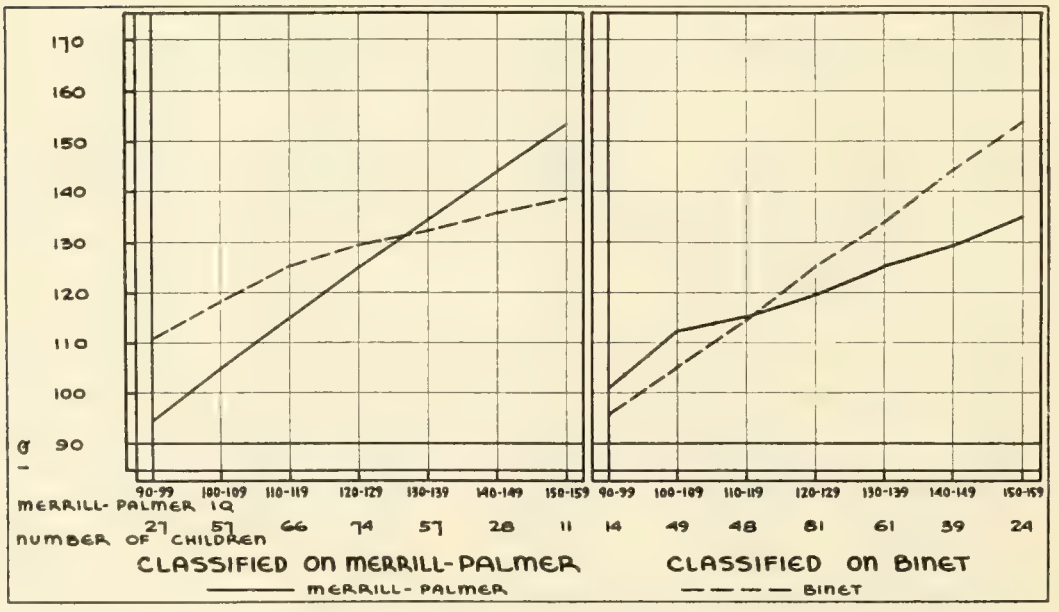


Figure 10. Correspondence Between Merrill-Palmer and Binet Means

9 and 10. Children with Merrill-Palmer quotients from 50 to 129 received higher Binet quotients than Merrill-Palmers, while children with Merrill-Palmer quotients of 130 and above received lower Binet quotients than Merrill-Palmer. If the two lowest divisions and the highest division are omitted because of small numbers of cases, it will be seen that the trend of divergence is very regular for the remaining groups. The greatest positive divergence (16.4 points) occurs at the level 90 to 99. Each succeeding division up to 130 to 139 shows a positive divergence of less amount than the preceding division. At 130 to 139 the divergence is the smallest (—2.0 points) of any level. Each succeeding division then diverges in greater amount in a negative direction.

There is the possibility that the reduction in divergence at the levels 50 to 89 is representative of the differences to be expected at these levels. This, however, needs to be checked with larger numbers of cases at these levels.

Since in most samplings a large proportion of the cases will be found in the average divisions, in general Binet intelligence quotients may be expected to run higher than Merrill-Palmer quotients. The amount of divergence will be a function of the relative proportions of the IQ levels represented.

When Grouped According to Binet IQ Level.—Table 35 and Figures 9 and 10 present similar material on divergence of Binet from Merrill-Palmer when the divisions are grouped according to Binet intelligence quotient. Children who received Binet quotients in the lower brackets (50 to 109) made higher scores on the Merrill-Palmer, while children who received very superior Binet quotients (120 to 179) lagged behind in Merrill-Palmer score. When the two lowest divisions are omitted because of small number of cases, the trend is regular, with the exception of less divergence at the 90 to 99 level than at the 100 to 109 level. The point of shift from negative to positive is the 110 to 119 level. The greatest divergence occurs at the "genius" levels, where the Merrill-Palmers are far behind the Binet quotients.

Considering Table 35, it will be seen that children who receive low Merrill-Palmer scores can be expected to receive higher Binet scores than their Merrill-Palmers, and children who receive low Binet scores can be expected to receive higher Merrill-Palmer scores than their Binets. Conversely, children who make very high Merrill-Palmer

scores can be expected to make somewhat lower Binet scores than their Merrill-Palmers, while children who make very high Binet scores can be expected to make Merrill-Palmer scores somewhat lower than their Binets.

At the same time, the mean on one of these tests rises consistently with the mean on the other test (Figure 10). The simultaneous rise in means indicates a communality of function in the two tests (This may be general intelligence or some attribute or attributes of intelligence in common.), while the divergence with IQ level suggests that the two tests are measuring somewhat different types of abilities. Perhaps this is the result of different emphasis on manipulatory and verbal abilities (31). If one can hazard a guess as to the primary abilities listed by Thurstone (32) which are predominant in the two tests, it might be suggested that the Binet tests include number facility, word fluency, verbal reasoning, and memory while the Merrill-Palmer test places more emphasis on visualizing and perceptual speed.

Correspondence Between Mental Ages

Table 36 shows the relationship between the two scales when the material is grouped according to mental age.

When Grouped According to Merrill-Palmer Mental Age.—Children with mental ages below 66 months on the Merrill-Palmer scale received mental ages on the Binet scale from 2.5 to 4.7 months higher. There was no particular tendency for the amount of divergence to be related to mental age level within these limits. Children who received very high Merrill-Palmer mental ages (66 to 78 months) were lower in Binet mental age.

When Grouped According to Binet Mental Age.—When the basis of grouping was Binet mental age, the correspondence between mental ages on the two tests appeared closer at the lower groupings than when the basis of grouping was the Merrill-Palmer scale. Children who made very high Binet mental ages (78 to 89 months) were decidedly lower in Merrill-Palmer mental age, the divergence being 8.6 months. The maximum Merrill-Palmer mental age is 78 months.

Changing Relationship After Preschool Attendance

The changing relationship between Merrill-Palmer IQ and Binet IQ after preschool attendance has been discussed in the section *Relation to Amount of Attendance Prior to First Test* (p. 70). It was pointed out that the correlation between the two tests for children new to pre-

school was .46 ± .06 (93 cases; see Table 34), while the initial Merrill-Palmer IQ of children new to preschool correlated with Binet IQ after 6 to 11 months of preschool attendance .24 ± .10 (45 cases). Similarly initial Binet of children new to preschool correlated with Merrill-Palmer (first test for these children) after 6 to 11 months of attendance .38 ± .07 (72 cases). Also simultaneous tests after 6 to 11 months of attendance correlated .39 ± .07 (72 cases). These results were interpreted as showing that preschool attendance may have lowered the relationship between the two tests. It appeared that the effect was greater on Binet abilities than on Merrill-Palmer abilities, although both were apparently affected.

These conclusions should be checked with further cases, since it will be seen later that initial Merrill-Palmers correlated with subsequent Binets at six-month intervals up to six years without reference to preschool attendance .38 to .46, and since the relationships between simultaneous tests in Table 34 tended to be higher than any of the above (.45 ± .05 to .63 ± .04).

Another line of attack on the problem of the effect of preschool experiences is the changing relationship between mental age and chronological age. The correlation between Merrill-Palmer mental age and chronological age for the 111 new entrants was .91 ± .01. For the seventy-five children who had their first Merrill-Palmer test after six or more months of preschool attendance the relationship was slightly reduced, the correlation being .86 ± .02.

The ninety-three new entrants who had Binet tests within one week of their Merrill-Palmer test showed a relationship of .90 ± .01 between Binet mental age and chronological age. This is almost identical with the Merrill-Palmer relationship for new entrants. However, there was greater reduction in Binet relationship after preschool attendance, the correlation dropping to .73 ± .04 for Binet as compared with .86 ± .02 for Merrill-Palmer of the same children.

Preschool Attendance, Months	Children	*r*	P.E.	Children	*r*	P.E.
0 to 5	111	.91	± .01	93	.90	± .01
6 or more	75	.86	± .02	74	.73	± .04

It appears that preschool attendance affected the relationship between Binet mental age and chronological age to a greater extent than it affected the relationship between Merrill-Palmer mental age and chronological age. A qualification to this conclusion is the fact that the longer attendance group had had some previous experience with

the Binet test, there being the possibility that experience with the test as well as preschool experiences may have contributed to displacement in relative position.

In both Merrill-Palmer and Binet abilities there was a reduction in the relationship between mental age and chronological age with preschool attendance.

Relation to Subsequent Binets

Of considerable interest is the problem of relationship of ability at the preschool ages to later intellectual development. Since the Merrill-Palmer scale is applicable only at the preschool ages, it becomes a serious question in longitudinal studies to decide what measures to use when the children have outgrown the scale. The present study throws light only on the relationship to subsequent Binets. Table 37 and Figure 11 present correlations between initial Merrill-Palmer intelligence quotients and Binet intelligence quotients at subsequent dates. For purposes of comparison, relationships between simultaneous early Binets and the subsequent Binets are presented for the same groups of children in Table 38. In Table 39, the two tests are combined and the mean IQ on the two tests related to the subsequent Binets. In Table 40 final Merrill-Palmer intelligence quotients are related to the subsequent Binets.

Relation of Merrill-Palmer to Subsequent Binets

Table 37 and Figure 11 show the relationship between initial Merrill-

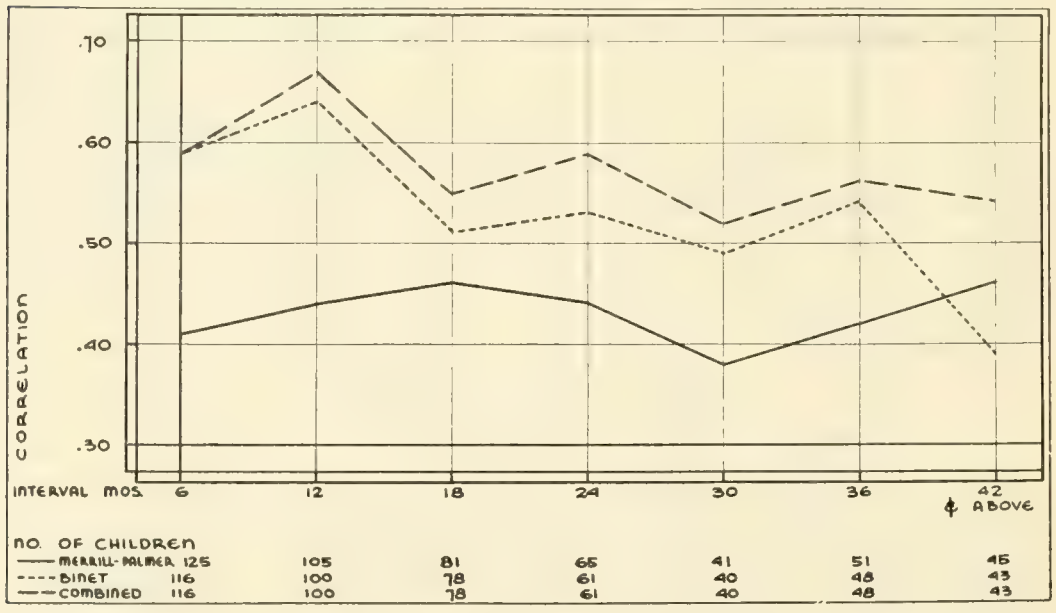


Figure 11. Relation to Subsequent Binets: Children Two to Five Years of Age on Initial Test

Table 37

Relation of Merrill-Palmer Intelligence Quotients (First Tests) to Subsequent Binets

Interval Between Merrill-Palmer and Subsequent Binet, Months	Chronological Age at Merrill-Palmer														
	18 to 29 Months			30 to 41 Months			42 to 53 Months			54 to 59 Months			18 to 59 Months		
	Children	r	P.E.	Children	r	P.E.	Children	r	P.E.	Children	r	P.E.	Children	r	P.E.
Six-Month Intervals															
3 to 8				46	.38 ± .08										
9 to 14				39	.44 ± .09								125	.41 ± .05	
15 to 20				34	.48 ± .09								105	.44 ± .06	
21 to 26				35	.50 ± .09								81	.46 ± .06	
27 to 32				27	.37 ± .10								65	.44 ± .07	
33 to 38				26	.52 ± .09								41	.38 ± .09	
39 to 70													51	.42 ± .08	
													45	.47 ± .08	
Twelve-Month Intervals															
6 to 17	39	.60 ± .07		81	.41 ± .06		75	.55 ± .05		32	.39 ± .10				
18 to 29	28	.37 ± .11		70	.49 ± .06		25	.34 ± .11							
30 to 41	24	.66 ± .10		49	.49 ± .07										

Table 38

Relation of Binet Intelligence Quotients to Subsequent Binets

Interval Between Tests, Months	Chronological Age on Initial Tests														
	18 to 29 Months			30 to 41 Months			42 to 53 Months			54 to 59 Months			18 to 59 Months		
	Children	r	P.E.	Children	r	P.E.	Children	r	P.E.	Children	r	P.E.	Children	r	P.E.
Six-Month Intervals															
3 to 8				44	.42 ± .09								116	.59 ± .04	
9 to 14				39	.61 ± .07								100	.64 ± .04	
15 to 20				33	.51 ± .09								78	.51 ± .06	
21 to 26				33	.56 ± .09								61	.53 ± .07	
27 to 32				26	.33 ± .14								40	.49 ± .08	
33 to 38				25	.41 ± .13								48	.53 ± .07	
39 to 70													43	.39 ± .09	
Twelve-Month Intervals															
6 to 17	37	.79 ± .04		78	.69 ± .05		71	.59 ± .05		26	.75 ± .07				
18 to 29	28	.57 ± .08		67	.54 ± .06		24	.49 ± .11							
30 to 41	24	.62 ± .10		48	.35 ± .09										

Table 39

Relation of Combined Merrill-Palmer and Binet Intelligence Quotients to Subsequent Binets

Interval Between Initial Tests and Subsequent Binets	Chronological Age on Initial Tests														
	18 to 29 Months			30 to 41 Months			42 to 53 Months			54 to 59 Months			18 to 59 Months		
	Children	r	P.E.	Children	r	P.E.	Children	r	P.E.	Children	r	P.E.	Children	r	P.E.
Six-Month Intervals															
3 to 8				44	.46 ± .08										
9 to 14				39	.66 ± .06								116	.59 ± .04	
15 to 20				33	.52 ± .09								100	.67 ± .03	
21 to 26				33	.61 ± .08								78	.55 ± .06	
27 to 32				26	.41 ± .10								61	.59 ± .06	
33 to 38				25	.56 ± .08								40	.52 ± .08	
39 to 70													48	.57 ± .06	
													43	.54 ± .08	
Twelve-Month Intervals															
6 to 17	57	.81 ± .04		79	.58 ± .05		71	.71 ± .04		26	.71 ± .07				
18 to 29	28	.58 ± .08		67	.58 ± .05		24	.55 ± .11							
30 to 41	34	.69 ± .08		48	.47 ± .07										

Table 40

Relation of Final Merrill-Palmer Intelligence Quotients to Subsequent Binets

Interval Between Last Merrill-Palmer and Subsequent Binet	Chronological Age on Merrill-Palmer, Months	Children	r	P.E.
Six-Month Intervals				
3 to 8	32 to 58	40	.36	± .09
9 to 14	32 to 58	32	.51	± .09
15 to 26	32 to 58	38	.24	± .10
27 to 69	32 to 58	29	.48	± .09
Twelve-Month Intervals				
6 to 17	32 to 58	71	.43	± .07
18 to 29	32 to 58	37	.25	± .10
30 to 69	32 to 58	25	.35	± .14

Palmer intelligence quotient and subsequent Binets at successive six-month intervals. One hundred twenty-five children received a Binet test six months (3 to 8 months) after their initial Merrill-Palmer test. The correlation between intelligence quotients at this interval was .41 ± .05. There was no apparent effect of length of interval on the relationship between Merrill-Palmer and Binet, since the correlations at successive six-month intervals up to 36 months varied only from .38 to .46. The number of cases varied from 41 to 125. The correlation for the forty-five cases with intervals longer than three years was .47 ± .08.

These relationships are not high for predictive purposes but are positive and significant.

Correlations were computed separately for children having their initial Merrill-Palmer at three years of age, this group having the largest number of cases of any age group. The correlations for children with initial tests at three years of age followed closely the pattern for the total group, the range being from .37 ± .10 to .52 ± .09. There was no definite tendency for the size of the correlations to change with increasing interval.

When the interval groupings were increased to a twelve-month range, it was possible to compare correlations for different age groups. These are shown in the lower portion of Table 37. No one age group was superior in correlation at all three intervals (12, 24, and 36 months). The highest relationships of the series were for children who were two years of age on their initial test. The relationship between Merrill-Palmer IQ at two years of age and Binet IQ one year later was .60 ± .07 and three years later .66 ± .10. However, the correlation

two years later dropped to .37 ± .11. There was no definite trend for the size of correlation to vary with length of interval within an age group, or for the size of correlation to vary with age at initial test when the interval was constant.

Relation of Binet to Subsequent Binets

For comparative purposes, the relationships between Binets simultaneous to the Merrill-Palmers and the subsequent Binets are presented in Table 38 and Figure 11. With one exception (interval 39 to 70 months) the correlations for the total group were higher throughout than for Merrill-Palmer. This should be expected since the relationships here are based on two measurements by the same scale. The correlations are not high for predictive purposes, ranging from .39 ± .09 to .64 ± .04. The median correlation was .53 as compared with .44 for Merrill-Palmers.

The correlations for the children having their initial test at three years of age followed fairly closely the pattern for the total group, ranging from .33 ± .14 to .61 ± .07. There was no particular trend for the correlations to vary in size with length of interval.

When the interval groupings comprised one year, there was a tendency for the size of correlation to be reduced with longer intervals. There was also with one exception a tendency for the correlations to be higher for children whose initial test was obtained at an earlier age. The highest correlations at each interval were for children initially tested at two years of age, these correlations ranging from .57 ± .08 to .79 ± .04.

Relation of Combined Merrill-Palmer and Binet Intelligence Quotients to Subsequent Binets

For each child an average IQ from the Merrill-Palmer and Binet tests was computed. This mean was then related to the subsequent Binets. (Table 39 and Figure 11) The combination resulted in correlations consistently somewhat higher than with Binet alone and considerably higher than with Merrill-Palmer alone. The range for the total group at six-month intervals was from .52 ± .08 to .67 ± .03. The median correlation was .57 as compared with .53 for Binet alone and .44 for Merrill-Palmer alone.

Thus combining Binet with Merrill-Palmer added a very considerable amount to the consistency of results of Merrill-Palmer with later Binet scores. However, adding Merrill-Palmer to the early Binet measurement did not increase the consistency of early Binet with later

Binet scores in amounts sufficient to justify the time and effort required in giving the Merrill-Palmer in most practical situations.

Unfortunately data are not available on the relationship of these early measures to later performance series such as the Pintner and Paterson series or the Arthur revision. It is conceivable that the Merrill-Palmer results will be more highly related than the Binet results to later intellectual ability as measured by performance scales.

The highest correlations between the combined measures and later Binets were for children initially tested at two years of age. At the one-year interval the correlation for the two-year-olds was .81 as compared with .71 at four years and at five years and .58 at three years. At the two-year interval, the correlation for two-year-olds was .58, the same as for three-year-olds, and the correlation for four-year-olds was .55. At the three-year interval the correlation for two-year-olds was .69 as compared with .47 for three-year-olds.

Relation of Combined Merrill-Palmer and Binet Intelligence Quotients to Subsequent Combined Quotients

It was possible to determine the relationship between combined quotients and subsequent combined quotients for two interval groups, six months and twelve months.

Interval, Months	Children	Correlation
3 to 8	54	.71 ± .05
9 to 14	33	.72 ± .06

The results were slightly higher than when either measure alone was related to subsequent Binets or when the combined measures were related to subsequent Binets. The correlations of .71 and .72 compare with .41 and .44 for initial Merrill-Palmer over the same intervals (Table 37), .59 and .64 for initial Binets over the same intervals (Table 38), and .59 and .67 for combined initial measures (Table 39).

Relation of Final Merrill-Palmers to Subsequent Binets

Since a number of the children had more than one Merrill-Palmer test, and since there were gains in IQ during preschool attendance, it seemed of interest to relate the final Merrill-Palmer IQ to subsequent Binets. There were forty children who had a Binet test six months (3 to 8 months) after their final Merrill-Palmer test, and thirty-two children who had a Binet test twelve months (9 to 14 months) after their final Merrill-Palmer. In order to have sufficient cases for correlation, it was necessary to combine the next two intervals making the interval range 15 to 26 months and the number of cases thirty-eight.

Twenty-nine children also received a Binet test 27 to 69 months after their final Merrill-Palmer.

There was no tendency for the size of the correlation to vary with length of interval, the range being from .24 to .51, the longest interval yielding a correlation of .48 and the shortest interval .36. The correlations were very similar to those obtained with initial Merrill-Palmers. The correlations for grosser interval groupings were similar also. All of the correlations were low, but positive.

Summary

Three hundred forty children were tested on both Merrill-Palmer and Binet scales at an interval not greater than one week in rotated order. The following conclusions were drawn from these data:

1. The correlations between Merrill-Palmer and simultaneous Binets were not high as between two different tests of intelligence. They were high enough (.45 to .63 for the different age groups) to indicate some communality between the two measures, but not high enough for predictive purposes. There was no tendency for the size of the correlation to increase or decrease proportionately to age from two years to five years.

2. When the children were classified according to Merrill-Palmer IQ level, it was found that there was a point (or level) at which the two scales gave similar mean IQ's. Below this point the divergence of Binet from Merrill-Palmer was positive and above this point negative, the divergence increasing in proportion to distance from the point. When grouped on Merrill-Palmer IQ the point of closest agreement was the 130 to 139 level, the positive divergence reaching an amount of 16.4 points at the 90 to 99 level and the negative divergence reaching 14.9 at the 150 to 159 level. When grouped on Binet IQ the point of closest agreement was the 110 to 119 level. Below this point the divergence of Binet from Merrill-Palmer was negative and above it positive, reaching an amount of 18.8 points at the 150 to 159 level.

a. Since in most samplings a large proportion of the cases will be found in the average divisions, in general Binet intelligence quotients may be expected to run higher than Merrill-Palmer quotients. The amount of divergence will be a function of the relative proportions represented at each IQ level.

b. Children who receive low Merrill-Palmer scores can be expected to receive higher Binet scores and children who receive low Binet scores can be expected to receive higher Merrill-Palmer scores. Conversely, children who receive very high Merrill-Palmer scores can be expected to make somewhat lower Binet scores, while children who receive very high Binet scores can be expected to make somewhat lower Merrill-Palmer scores. At the same time, the mean on one of these tests rises consistently with the mean on the other. The simultaneous rise in means indicates a communality of function (which may be general intelligence or some common attribute or attributes of intelligence). The divergence with IQ levels suggests that the two tests are measuring somewhat different types of abilities (perhaps different emphasis on manipulatory and verbal abilities), or when put into terms of Thurstone's (32) primary abilities, we may speculate that there is more emphasis in the Binet test on number facility, word fluency, verbal reasoning, and memory and more emphasis in the Merrill-Palmer test on visualizing and perceptual speed.

A large number of children received subsequent Stanford Binet tests at intervals ranging from six months to nearly six years. From these data the following conclusions were drawn:

1. There was no apparent effect of length of interval on the relationship between Merrill-Palmer and subsequent Binets, since the correlations at successive six-month intervals up to 36 months varied only from .38 to .46. For the forty-five cases with intervals greater than three years the correlation was .47 ± .08. These relationships are not high for predictive purposes.

2. There was no definite trend for the size of the correlation between Merrill-Palmer and subsequent Binets to vary with age at initial test. The correlations at twelve-month interval groupings varied from .34 to .66 ± .10, the latter being for two-year-old children retested on the Binet three years later.

3. Correlations between Binets simultaneous to the Merrill-Palmers and subsequent Binets were slightly higher throughout than for Merrill-Palmers. The correlations were not high for predictive purposes, ranging from .39 ± .09 to .64 ± .04 for six-month interval groupings. The median correlation was .53 as compared with .44 for initial Merrill-Palmers. When the interval range for a group was twelve months, the correlations ranged from .35 ± .09 (three years of age on initial test and three-year interval) to .79 ± .04 (two years of age on initial test and one-year interval). There was a tendency for the size of correlation to decrease with lengthening interval.

4. When for each child an average IQ on combined Merrill-Palmer and Binet tests was computed, the relationship to subsequent Binets rose. The combination resulted in correlations consistently somewhat higher than for Binet alone and considerably higher than with Merrill-Palmer alone. The range for six-month interval groupings was .52 ± .08 to .67 ± .03. The median correlation was .57 as compared with .53 for Binet alone and .44 with Merrill-Palmer alone. Adding Merrill-Palmer to early Binet did not increase the consistency of early Binet with later Binet in amounts sufficient to justify the time and effort required in giving the Merrill-Palmer in most practical situations. For separate age groups the highest correlation for the combined initial tests (.81 ± .04) was obtained for two-year-old children retested on the Binet one year later and the lowest (.47 ± .07) for three-year-old children retested three years later.

5. Correlations between final Merrill-Palmers and subsequent Binets were similar to those obtained for initial Merrill-Palmers.

6. The relationship between combined initial measures and combined subsequent measures was slightly higher than when related to subsequent Binet alone. The correlations for interval ranges of six months were .71 ± .05 for a six-month interval and .72 ± .06 for a twelve-month interval.

COMPARISON WITH RESULTS OF OTHER INVESTIGATIONS

There have been a number of significant studies of the Merrill-Palmer scale, and there has been some duplication of problems investigated. It is the purpose here to bring together the findings from the literature in order (1) to compare the results of the present study with them and (2) to synthesize the findings into general conclusions with respect to intellectual growth of young children. In the material which follows, the results of the various investigations are presented in rearranged form in order to facilitate comparisons.

Relation to Chronological Age

Mental Age

Correlations between mental age and chronological age have been reported by Bristol (4) and Goodenough (9). Bristol applied the Merrill-Palmer scale to 531 English children, ages twenty-four to seventy-one months. Goodenough's group was thirty-three children, ages twenty-seven to fifty-nine months, attending nursery school. The correlations obtained were .90 and .85 respectively (Table 41). The correlation obtained in this study (.81 ± .01) on first tests of 281 children, of whom 208 were preschool laboratory children, was somewhat lower, but the correlation based upon fall tests of the seventy-two preschool children who were retested in the spring was .91 ± .01,

Table 41
Relation to Chronological Age

Investigator	Children	Age Range, Months	r	P.E.
Mental Age				
Bristol (4)	531	24 to 71	.90	
Goodenough (9)	33	27 to 59	.85	
Wellman, first tests	281	20 to 62	.81	± .01
Wellman, fall tests	72	24 to 52	.91	± .01
Wellman, spring tests	72	30 to 58	.87	± .02
Wellman, first tests of new entrants to preschool	111	20 to 62	.91	± .01
Wellman, first tests of children with 6 to 11 months of preschool attendance	75	30 to 58	.86	± .02
Intelligence Quotient				
Goodenough (9)	33	27 to 59	.24	
Wellman, first tests	281	20 to 62	.12	± .04
Raw Score				
Stutsman (27)	631	18 to 77	.92	± .004
Kawin (15), first tests	169	23 to 57	.79	
Kawin (15), second tests	56	28 to 64	.66	
Wellman, first tests	281	20 to 62	.82	± .01

almost identical with Bristol's correlation. The correlation for the same children retested in the spring was slightly reduced (.87 ± .02), significant gains in IQ having taken place in the intervening period. When the first test group was restricted to new entrants to preschool (children who had had 0 to 3 months of attendance prior to their test), the correlation for 111 children was again .91 ± .01, while the relationship for the seventy-five children with 6 to 11 months prior preschool attendance was .86 ± .02.

Each of the studies indicates a very high correspondence between mental age and chronological age, in fact probably too high for adequate differentiation of mental age apart from chronological age.

Intelligence Quotients

Goodenough (9) reports a correlation of .24 between intelligence quotient and chronological age for the thirty-three nursery school children, indicating some tendency for the older children to possess higher IQ's. In this study the relationship based on first tests of 281 children was .12 ± .04, indicating but slight tendency for the intelligence quotients to vary consistently with age on first test.

Raw Scores

Correlations between chronological age and raw score are presented by Stutsman (27) and Kawin (15). Stutsman's group was composed of 631 children eighteen to seventy-seven months of age. Her correlation was .92 ± .004 (Table 41), indicating a very high relationship between raw score and age. Kawin's group of 169 children, ages twenty-three to fifty-seven months, showed considerably less relationship between raw score and age, her correlation being .79. For the fifty-six children having a retest, Kawin reports a still lower relationship, the correlation being reduced to .66. Apparently, some experiences intervening between first test and retest or some factors inherent in the test operated differentially at the different ages to produce a reduction in the relationship between chronological age and raw score. Possibly the group retested would have shown a lower relationship on first test than her total first test group.

The correlation between raw score and chronological age on first tests in this study, .82 ± .01, was very similar to that of Kawin's first test group and lower than Stutsman's figure.

Sex Differences in Percentile Rank

Stutsman (27) presents median percentile ranks for each sex in the different groups of children used in standardization of the Merrill-

Palmer scale. These are given in Table 42. In four of the five groups the median percentile rank was higher for girls than for boys, although in two groups the differences were negligible. The children in our group of new entrants to preschool (Tables 42 and 26) compare very closely in percentile rank with the Merrill-Palmer nursery school group, but the mean difference in favor of the girls is slightly greater in our group. The ratio of the difference in our group of new entrants was 1.1, or 86 chances in 100 of a true superiority of the girls. When the group compared was children who had attended preschool but who were receiving their first Merrill-Palmer test (Tables 42 and 26) the superi-

Table 42
Sex Differences in Percentile Rank

Source	Boys		Girls		Difference
	Children	Percentile Rank*	Children	Percentile Rank*	
Stutsman (27)					
Merrill-Palmer nursery school	72	74.0	59	75.3	1.3
Merrill-Palmer waiting list	93	51.9	96	52.7	0.6
Schools	47	40.8	47	57.5	16.7
Child care agencies	52	36.7	48	42.5	5.8
Health clinics	27	62.5	40	50.0	—12.5
Wellman					
0 to 3 months attendance prior to first test	50	73.3	61	78.1	4.8
6 to 11 months attendance prior to first test	47	78.2	28	86.2	8.0
Occupational group I	84	75.3	72	81.1	5.8
Occupational groups II to IV	25	80.1	27	78.6	— 1.5

*Median percentile rank in Stutsman, mean percentile rank in Wellman.

ority of the girls was even greater. The difference of 8.0 percentile ranks yielded a ratio of 1.8, or 96 chances in 100 of a true superiority of the girls.

When boys and girls whose fathers were in the professional classes (Occupational Group I) were compared, girls were 5.8 percentile ranks superior (Tables 42 and 29) while no differences were found between boys and girls whose fathers were in occupational groups II to IV. The number of cases in the lower classes was relatively small.

Other sex comparison will be presented in following tables in reference to changes on retests and to the effects of preschool attendance.

Changes According to Interval Between Tests

Changes according to interval between tests have been reported by

Table 43
Changes According to Interval Between Tests

Investigator	Interval, Months	Age Range on First Test of Series	Children	Mean on Initial Test	Mean on Retest	Difference	Ratio of Difference to Standard Deviation of Difference	Ratio of Difference to Probable Error of Difference	r	P.E.
Intelligence Quotients										
Driscoll (7)	6	24 to 42 Months	80	108.6	114.6	6.0	4.0		.62 ±	.05
Driscoll (7)	12	24 to 42 Months	40	105.4	113.2	7.8	2.95		.47 ±	.08
Wellman	5 to 7	20 to 62 Months	127	118.4	125.6	7.2	3.6		.63 ±	.04
Wellman	11 to 13	20 to 62 Months	53	116.2	128.4	12.2	4.2		.52 ±	.07
Sigma Scores										
Stutsman (28)	2	2 to 5 Years	77						.72 ±	.04
Stutsman (28), first to second test	6 to 9	2 to 5 Years	207						.59 ±	.03
Stutsman (28), second to third test	6 to 9	2 to 5 Years							.71 ±	.03
Kawin (15), first to second test	2 to 29	23 to 57 Months	169	.065	.379	.314		8.4	.59 ±	.03
Kawin (15), second to third test	2 to 19	28 to 64 Months	56	.552	.638	.086		1.0	.57 ±	.06
Kawin (15), first to third test		23 to 57 Months	56	— .096	.638	.734		9.4	.49 ±	.07
Wellman	¼	28 to 60 Months	44	— .14	.40	.54			.92 ±	.02
Wellman	5 to 7	20 to 62 Months	127	.93	1.12	.19	2.1		.54 ±	.05
Wellman, first to second test	5 to 7	20 to 62 Months	68	.77	1.12	.35	2.6		.55 ±	.06
Wellman, second to third test	5 to 7	20 to 62 Months	41	1.08	.99	— .09	.6		.58 ±	.07
Wellman	11 to 13	20 to 62 Months	53	.96	1.11	.15	1.1		.50 ±	.07
Wellman	15 to 20	20 to 62 Months	24	.88	1.22	.34	1.4		.47 ±	.11

Raw Scores									
Stutsman (28)	1 day to 1 month	3 to 5 Years	46						.85 ± .03
Stutsman (28)	2 or less	2 to 5 Years	81						.95 ± .01
Stutsman (28)	6 or less	2 to 5 Years	64						.91 ± .01
Stutsman (28)	1 to 11	3 Years	56						.72 ± .05
Kawin (15), first to second test	2 to 29	23 to 57 Months	169	34.5	56.7	22.2		34.9	.72 ± .02*
Kawin (15), second to third test	2 to 19	28 to 64 Months	56	49.8	66.8	17.0		18.4	.73 ± .04
Kawin (15), first to third test		23 to 57 Months	56	26.8	66.8	40.0		43.8	.67 ± .04

*When chronological age was held constant, the correlation was .59 ± .03.

Driscoll (7), Stutsman (28), and Kawin (15). The results of our study are compared with theirs in Table 43.

Intelligence Quotients

Driscoll's (7) results and our results are in close agreement, in that significant gains in intelligence quotient were found over six- and twelve-month periods. Our group tested initially higher than Driscoll's group, but made approximately the same amount of gain over a six-month period (6.0 and 7.2 points respectively). Over a twelve-month period the gain of our group was larger than for Driscoll's group, the gain being 12.2 points in our group as compared with 7.8 points in Driscoll's group. The correlations are strikingly similar, each investigator showing a drop in correlation with increased interval between tests (.62 and .63 at six months, .47 and .52 at twelve months).

Sigma Scores

Kawin's (15) children gained significantly in sigma score on retest from their first to their second test. From second to third tests the gain was not significant. Our groups of children retested at six-, twelve-, and eighteen-month intervals tested initially higher than Kawin's group and the gains were not entirely significant. Our results agree with Kawin's in showing no gain from second to third tests. In our group the explanation for the differences appears to lie in the amount of intervening preschool attendance (see Table 10 and *Effects of Preschool Attendance* (p. 52). Our group retested at an interval of one week started at a lower level and made a greater gain than our other groups. This gain was attributed to practice effects (see *Change on Retests* (p. 31).

Correlations of sigma scores reported by the different investigators are in quite close agreement, with the exception that Stutsman's (28) figure of .71 ± .03 between second and third tests is higher than Kawin's figure of .57 ± .06 between second and third tests. The age ranges are comparable. Stutsman's interval range is much more restricted, being limited to 6 to 9 months, while Kawin's interval range is 2 to 19 months. Our correlation between second and third tests at an interval closely comparable to Stutsman's (5 to 7 months in our group) was .58 ± .07, a figure which is strikingly in agreement with Kawin's results and lower than Stutsman's. Stutsman's correlation between first and second tests agrees very closely with the results reported by Kawin and our study. The respective correlations were .59, .59, and .55.

The highest correlation on sigma scores in the table, .92 ± .02, is

that obtained in our study, on retests at one week. Stutsman's correlation of .72 ± .04 at a two-month interval is the next highest figure and the next shortest interval. Our results show some further decrease in the correlation with length of interval from six to eighteen months, the correlation at six months being .54, at twelve months .50, and at eighteen months .47. Kawin's results also indicate reduction in correlation with increasing interval, since her correlation between first and third tests was .49 as compared with .59 between first and second tests. Stutsman's correlation between second and third tests (.71 ± .03), however, was as high as her correlation at two months.

Kawin's (15) correlations between Binet IQ's on retests were higher than her correlations between sigma scores. For 114 children retested at an interval of 10.6 months, the correlation between first and second test Binet IQ's was .75 ± .03 as compared with .59 ± .03 for the 169 children on sigma score retests at 9.2 months. For thirty-nine children the Binet correlation between second and third tests was .84 ± .03, as compared with .57 ± .06 for sigma scores on second and third tests for fifty-six children retested at an interval of 8.7 months. From first to third tests, the correlation on Binet was .75 ± .05 (39 cases) as compared with .49 ± .07 on sigma scores (56 cases).

Our correlations for Binets were lower than Kawin's. For the total group (N = 119) with a six-month interval the correlation was .64 ± .04. For the sixty-three children from first to second test the correlation was .67 ± .04, for the thirty-eight children from second to third test .61 ± .07, and for the thirty-five children from first to third test .69 ± .06. For the fifty-one children with 11 to 13 months between tests the correlation was .66 ± .05. These were all higher than comparable figures for sigma scores.

Raw Scores

Correlations between raw scores on retests are presented by Stutsman and Kawin. It is to be expected that these correlations will be higher than for intelligence quotients or sigma scores, since age is a factor in determining relative ranking in raw scores. Stutsman (28) has pointed out that narrowing the age range of the group reduces the correlation. The restriction of one of her groups to a one-year age range apparently explains the correlation of .72 as compared with the higher values (.85 to .95) when the age range was wider. Kawin reported a correlation of .72 between raw scores on first and second tests, which was reduced to .59 when chronological age was held constant.

The latter value compares very closely with the values obtained for sigma scores by Stutsman, Kawin, and our study, and for intelligence quotients obtained by Driscoll and our study.

Age Differences on Retests, Compared With Binet

Comparison of gains over a six-month interval by different age groups has been made by Driscoll (7). The children in her study were examined in the nursery school and clinic of the Child Development Institute of Teachers College, Columbia University. The two sources were combined in the treatment of data. Half-year age groups were compared, the exact year or exact half-year forming the midpoint of a group. For example, the age 24 months included a range from 21 months to 26.9 months. The intelligence quotients used were converted measures. The mental age obtained by a child was transferred into a standard deviation value which in turn was transferred into its appropriate IQ value. Driscoll states that the difference beween the IQ obtained in this way and the IQ obtained directly by dividing mental age by chronological age ranged from 1 to 5 points.

From Table 44 it will be seen that significant gains were made over a six-month interval at ages 30 and 36 months, but that no gain was made at 24 months. In our group, two-year-old children gained as much as three- and four-year-old children. The gains were quite similar to Driscoll's except at this one age level.

Driscoll's correlations indicate no particular trend for consistency on retest to increase or decrease with age. The correlation at 42 months was lower than at the other ages, but the number of cases was small.

There was an increase in the mean IQ with age. In Driscoll's data the means varied from 106.8 at 24 months of age to 122.9 at 42 months. Our data showed similar differences: 108.9 at two years, 117.1 at three years, and 123.5 at four years. Binet means were higher throughout.

Driscoll's correlation (.62 ± .05) for the entire age range, 24 to 42 months, on eighty cases, is almost identical with our correlation (.63 ± .04) on 127 cases 20 to 62 months of age.

Binet correlations tended to run slightly higher in each of Driscoll's groups than Merrill-Palmer correlations and there was a rise in Binet correlation with increasing age. However, when the total age range of 24 to 42 months was used the correlation for 109 cases was almost identical (.61 ± .04) with the Merrill-Palmer correlation. Our Binet correlation for 119 children over a six-month interval was the same (.63 ± .04) as the Merrill-Palmer correlation.

Table 44

Changes by Age Groups on Retests, Compared with Binet

Investigator	Chronological Age, Months	Interval, Months	Children	Merrill-Palmer						Binet						
				Intelligence Quotient		Change in IQ	Ratio of Difference to Standard Deviation of Difference	r	P.E.	Children	Intelligence Quotient		Change in IQ	Ratio of Difference to Standard Deviation of Difference	r	P.E.
				Earlier Test	Later Test						Earlier Test	Later Test				
Changes on Retest																
Driscoll (7)	24	6	37	106.8	107.8	1.0		.56 ± .08		37	120.4	122.4	2.0		.58 ± .07	
Driscoll (7)	30	6	47	107.0	114.3	7.3	3.8	.54 ± .07		42	119.5	127.6	8.1		.64 ± .06	
Driscoll (7)	36	6	33	111.9	122.3	10.4	4.5	.60 ± .08		43	122.1	127.4	5.3		.73 ± .05	
Driscoll (7)	42	6	16	122.9	128.4	5.5		.39 ± .14		12	129.7	128.8	— .9		.80 ± .07	
Driscoll (7)	24 to 42	6	80	108.6	114.6	6.0	4.0	.62 ± .05		109	117.6	121.9	4.3	3.2	.61 ± .04	
Wellman	18 to 29	5 to 7	17	108.9	118.8	9.9										
Wellman	30 to 41	5 to 7	64	117.1	124.4	7.3										
Wellman	42 to 53	5 to 7	45	123.5	129.7	6.2										
Wellman	20 to 62	5 to 7	127	118.4	125.6	7.2	3.6	.63 ± .04		119	129.7	131.6	1.9	.8	.63 ± .04	
Driscoll (7)		12	40	105.4	113.2	7.8	3.0	.47 ± .08		91	117.2	122.9	5.7	3.7	.60 ± .05	
Wellman	20 to 62	11 to 13	53	116.2	128.4	12.2	4.2	.52 ± .07		51	130.2	132.6	2.4	.7	.63 ± .06	
Driscoll (7)		18								38	117.9	125.5	7.6	2.7	.56 ± .08	
Wellman	20 to 62	15 to 20	24	111.8	132.9	21.1	4.6	.50 ± .11		24	127.2	126.4	.8	.2		
Fall to Spring																
Wellman	20 to 62	5 to 7	72	118.1	127.2	9.1	3.4	.60 ± .05		66	127.1	131.8	4.7	1.7	.61 ± .05	
Spring to Fall																
Wellman	20 to 62	5 to 7	46	120.7	124.4	3.7	1.1	.73 ± .05		43	134.9	130.4	—4.5	3.7	.65 ± .06	

It will be noted that Driscoll's correlations for the Merrill-Palmer correspond more closely to our correlation covering a period of preschool attendance (.60 ± .05 for fall to spring tests) than to our correlation covering a period of nonattendance (.73 ± .05 for spring to fall tests).

When the interval between tests was a year, the correlation was higher for Binet than for Merrill-Palmer. Driscoll's and our results are in close agreement (.47 and .52 for Merrill-Palmer, .60 and .63 for Binet).

Sex Differences on Retests, Compared With Binet

The percentages of each sex who gained on retest at six months and at twelve months in Driscoll's (7) group and in our groups are compared in Table 45. More of our boys than Driscoll's (71 per cent versus 56 per cent) gained over a six-month period and more of our girls than Driscoll's gained over a twelve-month period (90 per cent versus 64 per cent). Otherwise the percentages were very similar. The results do not permit any generalization as to sex differences in number gaining on retest.

The two studies agree in showing a smaller percentage gaining on the Binet test than on the Merrill-Palmer. The percentages gaining on Binet retests ranged from 43 to 71. In both studies when the Merrill-Palmer gains showed differences between the sexes in percentage gaining, there was the same direction of difference on the Binet. Since in our group, the same children were involved in Binet and Merrill-Palmer, it appears that whatever induces gains on one scale also affects the other scale.

The amount of gain and loss as related to sex differed in the two studies. In both interval groups Driscoll's boys made larger median gains and smaller median losses than the girls. Our results were conflicting. On Binet test all groups of boys made larger median gains than girls, although the difference was only .5 point in one of our groups. The median gains of the boys ranged from 11.0 to 20.0, and those of the girls from 7.0 to 11.0. Results on median losses were conflicting. The range of median losses of boys was 5.0 to 10.5 and of girls 6.0 to 9.8.

Gains on Retests According to IQ Level

In Table 46 are compared the percentages of children gaining and losing and the median gains and losses on Merrill-Palmer and Binet retests at a six-month interval. In both Driscoll's (7) and our groups there was an increase in the percentage gaining with decrease in initial

Table 45
Sex Differences in Changes on Retests, Compared with Binet

Investigator	Sex	Interval, Months	Chronological Age, Months	Merrill-Palmer				Binet			
				Children	Per cent Gaining	Median of Gains	Median of Losses	Children	Per Cent Gaining	Median of Gains	Median of Losses
Retests											
Driscoll (7)	Boys	6	24 to 42	39	56	14.8	6.3	53	43	20.0	6.5
Driscoll	Girls	6	24 to 42	41	73	13.0	9.8	56	61	7.0	6.5
Wellman	Boys	5 to 7	20 to 62	70	71	11.0	6.0	68	57	11.0	10.5
Wellman	Girls	5 to 7	20 to 62	57	72	12.0	9.0	51	59	10.5	8.0
Driscoll (7)	Boys	12	24 to 42	15	80	16.5	5.0	38	71	15.3	6.0
Driscoll	Girls	12	24 to 42	25	64	11.5	18.0	53	51	8.4	9.8
Wellman	Boys	11 to 13	20 to 62	35	75	15.0	8.0	31	45	15.0	5.0
Wellman	Girls	11 to 13	20 to 62	33	90	12.0	3.5	20	65	11.0	6.0
Fall to Spring Retests											
Wellman	Boys	6	20 to 62	39	74	13.0	7.5	37	65	11.0	14.0
Wellman	Girls	6	20 to 62	33	76	11.0	6.0	29	76	7.0	7.0
Spring to Fall Retests											
Wellman	Boys	6	20 to 62	27	74	9.5	5.0	25	44	6.0	8.5
Wellman	Girls	6	20 to 62	19	47	11.0	13.0	18	28	14.0	10.0

IQ classification, with the exception of a reversal of position of Driscoll's superior and normal groups. Binet changes showed the same tendency, except for an additional reversal of position between superior and normal in our group.

Median gains in Driscoll's group showed a tendency to increase with decrease in initial IQ. This tendency was not evident in our data. Median losses were irregular in amount at the different IQ levels on the Merrill-Palmer. The Binet showed a regular increase in median gain from higher to lower classifications. In Driscoll's group the median gains on Binet varied from 4.0 at the genius level to 19.0 at the normal level. In our group the median gains varied from 7.0 to 13.0.

Table 46
Gains Over Interval of Six Months According to IQ Level

Terman Classification	Children		Per Cent Gaining		Per Cent Losing		Per Cent Same		Median of Gains		Median of Losses	
	Driscoll (7)	Wellman	Driscoll (7)	Wellman	Driscoll (7)	Wellman	Driscoll (7)	Wellman	Driscoll (7)	Wellman	Driscoll (7)	Wellman
Merrill-Palmer												
Genius	0	11		9		91		0		7.0		12.0
Very Superior	18	48	50	71	50	25	0	4	5.8	12.5	8.0	5.5
Superior	14	28	71	79	29	18	0	4	15.0	11.0	13.5	3.0
Normal	39	39	64	85	31	13	5	3	12.3	10.0	6.8	6.0
Dull	7	1	86	100	14	0	0	0	16.3		4.0	
Borderline	2	0	100		0		0		25.5		0.0	
Binet												
Genius	7	33	29	33	57	67	14	0	4.0	7.0	12.5	13.0
Very Superior	43	51	56	59	39	39	4	2	9.8	7.5	5.7	8.0
Superior	26	17	38	88	58	6	4	6	15.0	11.0	6.0	4.0
Normal	32	18	66	72	28	22	6	6	19.0	13.0	2.0	2.5

Likewise there was a regular decrease in the median losses on Binet from high to lower classifications. In Driscoll's group the median losses on Binet varied from 12.5 at the genius level to 2.0 at the normal level. In our group the median losses varied from 13.0 to 2.5 points.

Changes During Periods of Preschool Attendance and Nonattendance

The effects of preschool attendance upon changes in Merrill-Palmer intelligence have been studied by Kawin and Hoefer (16), Barrett and Koch (2), and Waring (36). The former two used the method of

comparison by control groups. Kawin and Hoefer carefully matched twenty-two pairs of children applying for admission to two nursery schools, the Winnetka school whose entrants were of very high socio-economic status and the Franklin school of lower status. The children who were admitted to these two nursery schools were matched with other nonadmitted children on the basis of age, mental age, sex, and economic status. Just prior to entering nursery school, the preschool group was 0.2 months higher than the control group in mental age and 3.5 percentile ranks superior. After 7.8 months, the two groups were still almost identical in ability, the difference being 0.1 month in mental age and 2.1 percentile ranks (Table 47). The preschool group thus did not show greater growth associated with their nursery school experience. Both groups had gained large amounts in percentile rank, 24.6 and 26.0 respectively. Their mental age increase (11.4 and 11.5 months) was greater than the chronological age difference. Barrett and Koch (2) studied two orphanage groups matched on sex, age, initial Merrill-Palmer score, and length of institutional experience. After an interval of six to nine months in which one group of seventeen children attended nursery school, but otherwise led the same life as the control children, retests were made. The preschool group made a very large gain, 20.9 points in IQ, while the control group gained 5.1 points. It appeared that the nursery school experiences had definitely resulted in accelerating the IQ of these children.

The gain made by Barrett and Koch's nursery school children was considerably larger than the gain made by the preschool children in our study. However, the orphanage children were considerably lower in initial ability, their mean IQ being 91.7, while the mean IQ of the Iowa City children was 118.1. Out of the Iowa City group twenty-three children who tested 90 to 109 on their fall test, with a mean of 100.6, gained 14.0 points from fall to spring. The 90 to 99 portion of the Iowa City fall group was composed of nine children who gained 15.3 points; the 100 to 109 portion was composed of fourteen children gaining 13.2 points. The fourteen children testing 90 to 109 in the spring, with a mean of 102.1, gained 5.1 points from spring to fall, the same amount of gain as Barrett and Koch's control group. Thus the two studies agree not only in trend, but fairly closely in amount of gain when initial ability is taken into consideration.

From the mean chronological ages and mean mental ages given by Kawin and Hoefer (16), intelligence quotients were computed in order

Table 47
Changes During Periods of Preschool Attendance and Nonattendance

Investigator	Basis of Matching Groups	Group	Children	Prior to Attendance			After Attendance			Interval Months	Change		
				Mental Age, Months	Intelligence Quotient	Percentile Rank	Mental Age, Months	Intelligence Quotient	Percentile Rank		Mental Age, Months	Intelligence Quotient	Percentile Rank
Kawin and Hoefer (16)	Age, mental age, sex, economic status	Preschool	22	30.5	96.5*	43.0	41.9	109.1*	67.6	7.8	11.4	12.6*	24.6
		Control	22	30.3	95.6*	39.5	41.8	108.6*	65.5	7.8	11.5	13.0*	26.0
Barrett and Koch (2)	Sex, score, age, institutional experience	Orphanage, preschool	17		91.7			112.6		6 to 9		20.9	
		Orphanage, control	17		92.6			97.7		6 to 9		5.1	
Waring (36)	Nursery school children	Preschool, fall to spring	103							5	10		
		Preschool, spring to fall	21							8**	10		
Wellman	Preschool children	Preschool, fall to spring	72		118.1	79.0		127.2	87.2	5.6		9.1	8.2
		Preschool, spring to fall	46		120.7	82.7		124.4	84.8	6.4		3.7	2.1
Wellman	Identical preschool children	Preschool, fall to spring	36		114.0	72.2		122.7	84.6	5.7		8.7	12.4
		Preschool, spring to fall	36		118.8	81.1		123.0	84.0	6.4		4.2	2.9

*Computed from mean mental ages and chronological ages given by authors.
**Approximately half of these months involved no nursery school attendance.

to make a more direct comparison with Barrett and Koch's results and the results of this study. It will be seen that the mean initial intelligence quotients fell within the 90 to 109 level, and that the gains made by the preschool group correspond rather closely with the gains of the preschool children of that level of initial ability in our study, while the gain of Kawin and Hoefer's control group far exceeded the summer gains of Iowa City children and of the control orphanage group of Barrett and Koch.

Percentile rank comparisons reveal similar agreement of Kawin and Hoefer's preschool group with Iowa City children and disagreement of their control group with summer changes of Iowa City children. The gain in percentile rank from fall to spring for the twenty-three children testing 90 to 109 IQ was 22.0 ranks as compared with 24.6 for Kawin and Hoefer's preschool group, while the spring to fall change for the fourteen children similarly classified was 9.7 percentile ranks as compared with 26.0 for Kawin and Hoefer's control children.

Waring's study (36), while based upon a larger number of children, is reported very briefly. In all probability, the initial level of the group corresponded more closely with the total Iowa City group than was the case in the other two studies, since the children were in attendance at the Cornell University nursery school. In five months of preschool attendance, the children accomplished as much gain in mental age as they accomplished in eight months of vacation time.

Thus all of the studies that have made this type of analysis agree in gains of the children while attending preschool. The amount as well as the fact of gain is in close agreement, considering the possible variations in preschool environment. The different preschools appeared to be about equally effective. All of the studies agree also in the fact of gain of the control groups, but the one study by Kawin and Hoefer deviates from the others in the amount of gain.

Relation to Number of Days of Preschool Attendance Prior to Test

Only one other study has reported on the relationship of the child's intelligence to the number of days of preschool attendance prior to the test. Goodenough (9) included correlations between number of days of attendance and Merrill-Palmer ability and Minnesota Preschool Scale ability for thirty-three children, as a rather incidental part of a study on interrelationships in behavior. The correlations with

Table 48
Relation to Number of Days of Preschool Attendance Prior to Test

Investigator	Children	Sex	Age Range, Months	Days Attendance	Correlations		
					Mental Age	Mental Age with Chronological Age Constant	Intelligence Quotient
Merrill-Palmer Scale							
Goodenough (9)	33		27 to 59		.74	.42	.46
Wellman	95	Boys	20 to 62	1 to 280			.13 ± .07
Wellman	89	Girls	20 to 62	1 to 280			.41 ± .06
Minnesota Preschool Scale							
Goodenough (9)	33		27 to 59		.73	.44	.30

mental age were .74 and .73 respectively for the two scales (Table 48), indicating considerable tendency for the children with the longer amounts of attendance to possess the higher mental ages. When chronological age was held constant, the partial correlations of mental age with number of days of preschool attendance were .42 and .44. Correlations of IQ with days attendance were .46 and .30. The two scales thus gave very comparable relationships, IQ on the Minnesota Scale being less related to number of days attendance than IQ on the Merrill-Palmer scale.

Our result for eighty-nine girls was in close agreement with Goodenough's correlation, but our correlation for ninety-five boys was only .13 ± .07, indicating no particular tendency for the intelligence quotient of boys to be related to number of days they had attended preschool. When sigma score and percentile ranks instead of IQ were correlated, no relationship was found for the girls, the correlations being .10 ± .07 for sigma score and .18 ± .07 for percentile rank.

The amount of attendance in our study varied up to 280 days and was fairly equally distributed for boys and girls. Forty-eight per cent of the boys and 66 per cent of the girls had attended less than forty-one days. Forty-five per cent of the boys and 30 per cent of the girls had attended more than eighty days (equivalent of four preschool months).

In view of the conflicting correlations for girls when different methods of scoring were used, conclusions regarding the size of relationship between amount of attendance and Merrill-Palmer ability should be tentative.

Correspondence Between Merrill-Palmer and Binet

Relation to Binet Means

Mental Age.—Comparisons between Merrill-Palmer and Binet mental ages have been made by Cushing (6) and Rust (25). Their results (Table 49) are in agreement with ours in finding a higher mental age on the Binet than Merrill-Palmer. The difference in Rust's (25) group was shown to be statistically significant.

Correlations between Merrill-Palmer and Binet mental ages presented by Kawin (15) and Stutsman (27) agree very closely, all being .78 or .79. Rust's (25) correlation of .63 is lower, but her age group was much more restricted. Kuenzel's (17) correlation of .94 for children in the Mooseheart Baby village was very much higher than that found by any other investigator.

Intelligence Quotients.—Driscoll (7), Jersild and Markey (14), Nelson (23), and Rust (25) have compared Merrill-Palmer and Binet IQ's. All agree in finding higher Binet than Merrill-Palmer IQ's (Table 49). Nelson's group showed less superiority than the other groups of the same level and age.

Correlations reported by Driscoll (7), Nelson (23), and Rust (25) ranged from .55 ± .06 to .76 ± .05. Our correlations tend to run somewhat lower although overlapping (.45 ± .05 to .63 ± .04).

By IQ Levels.—While our children showed a definite trend for the difference between Merrill-Palmer and Binet IQ to vary with IQ level, Rust's (25) group did not show this trend, except that children with Binet IQ's of 100 to 120 were less superior over their Merrill-Palmers (5.2 points) than children with Binet IQ's above 120 (superiority 12.5 to 15.6 points):

Binet IQ Level	Children	Difference*
Rust, Ages 33 to 39 Months		
141 to 175	27	12.5
131 to 140	33	13.7
121 to 130	14	15.6
100 to 120	26	5.2
Wellman		
140 to 179	72	19.1
130 to 139	61	8.7
120 to 129	81	5.5
100 to 119	97	— 4.0

*Binet is higher

Our group testing at 100 to 119 was 4.0 points inferior in Binet, while Rust's children at this level were 5.2 points superior.

Rust used the Kuhlmann-Binet entirely with a restricted age range of 33 to 39 months, while our group was a mixture of Kuhlmann- and

Table 49
Relation to Binet

Investigator	Children	Chronological Age, Months	Merrill-Palmer Mental Age, Months	Merrill-Palmer Score	Binet Score	Difference	Ratio of Difference to Probable Error of Difference	r	P.E.
			Mental Age						
Cushing (6)	49	23 to 60		42.0	49.1	— 7.1			
Kawin (15)	55							.78 ±	.04
Kuenzel (17)	64	30 to 72						.94 ±	.01
Rust (25)	100	33 to 39		43.2	48.0	— 4.8	5.2	.63 ±	.04
Stutsman (27)	159	36 to 72						.79 ±	.02
Stutsman (27), feeble-minded group	29	54 to 150						.79 ±	.05
Stutsman (27), nursery school	115							.78 ±	.03
Wellman	21		18 to 29	25.9	29.6	— 3.7			
Wellman	80		30 to 41	35.4	39.3	— 3.9			
Wellman	74		42 to 53	46.7	51.4	— 4.7			
Wellman	81		54 to 65	57.9	60.4	— 2.5			
Wellman	52		66 to 77	72.6	69.2	3.4			
Wellman	19		78	78.0	73.1	4.9			

		Intelligence Quotient					
Driscoll (7)	53	21 to 26.9	108.4	120.3	—11.9	.63	± .05
Driscoll (7)	71	27 to 32.9	106.3	120.2	—13.9	.64	± .05
Driscoll (7)	85	33 to 38.9	111.5	127.7	—16.2	.70	± .04
Driscoll (7)	54	39 to 44.9	118.5	129.8	—11.3	.55	± .06
Driscoll (7)	30	45 to 50.9	117.3	124.0	—16.7	.76	± .05
Jersild and Markey (14), nursery school	17						
	15	22 to 36	112.0	136.0	—24.0		
Jersild and Markey (14), nursery school	19	33 to 50	125.0	134.0	— 9.0		
	102	33 to 39	101.9	107.6	— 5.7	.75	
Nelson (23)	100	33 to 39	119.1	132.7	—13.6	.60	± .04
Rust (25)	32	18 to 29	109.5	121.6	—12.1	.50	± .09
Wellman	126	30 to 41	114.6	123.3	— 8.7	.63	± .04
Wellman	125	42 to 53	125.3	128.8	— 3.5	.45	± .05
Wellman	57	54 to 65	124.6	122.3	2.3	.54	± .06
Wellman	340	18 to 65	119.7	125.0	— 5.3	.52	± .03
Wellman (0 to 3 months attendance prior to test)	93	20 to 59				.46	± .06

9.5

Stanford-Binets and included an age range of two to five years. However, that does not seem to explain the differences, since our trend was so consistently regular. Driscoll's (7) comparisons with Kuhlmann-Binet showed about the same percentage of children at 30 and 36 months making higher scores on the Kuhlmann as at the other ages. Her figures were:

Chronological Age, Months	Children	Per Cent Lower on Merrill-Palmer
24	53	76
30	71	87
36	85	86
42	54	80
48	30	77

Possibly the environments of Rust's children emphasized the types of abilities measured by the Binet test more relatively to Merrill-Palmer abilities than was true in our group, except for those children in our group who were extremely high in Binet ability.

The percentage of children at the different IQ levels for whom the Merrill-Palmer was lower was computed by Driscoll (7), for three-year-old children. The comparisons are presented in Table 50. Driscoll's results and ours are in close agreement in showing a regular trend for the percentage to rise with successively higher IQ classification. Driscoll's percentages run considerably higher than ours at each classification, but the trend is the same.

The median difference for children whose Merrill-Palmer was lower than their Binet increased regularly with increase in IQ in both Driscoll's and our groups. Children of average Binet IQ who were lower

Table 50
Percentage of Children Classified According to Binet IQ Who Differed in Merrill-Palmer IQ

Terman Classification	Children		Percentage of Children for Whom Merrill-Palmer IQ Was						Median Difference in IQ of Children Who Were			
			Higher		Lower		Same		Higher		Lower	
	Driscoll* (7)	Wellman	Driscoll (7)	Wellman	Driscoll (7)	Wellman	Driscoll (7)	Wellman	Driscoll (7)	Wellman	Driscoll (7)	Wellman
Genius	22	72	0	15	100	83	0	1	0.0	5.0	24.5	24.5
Very superior	34	142	9	36	91	62	0	2	6.5	7.0	20.8	14.5
Superior	16	48	7	50	87	46	7	4	13.0	10.5	13.5	9.5
Normal	10	63	10	57	60	37	30	6	9.0	17.0	8.5	9.0
Dull	3	9	100	78	0	11	0	11	4.5	10.0		10.0

*Three-year-old children.

in Merrill-Palmer IQ were 8.5 and 9.0 points lower in the two studies, while children of genius Binet IQ who were lower in Merrill-Palmer IQ were 24.5 points lower in both studies. In the intervening classifications Driscoll's groups showed larger differences than our groups.

Driscoll's study supports ours in showing an increasing divergence between Merrill-Palmer and Binet IQ's, both in percentage of children and in amounts of divergence, with increase in Binet IQ. As the percentage of children increases, the median difference increases also. Rust's (25) findings are not in agreement on the correspondence between the two tests.

Relation to Subsequent Binets

Driscoll (7) has presented comparisons between Merrill-Palmer and Kuhlmann tests as related to subsequent Stanford-Binets. Two groups were compared, one being two and three years of age on initial test and the other three and four years of age and both receiving the Stanford-Binet at five years of age or older (Table 51). Our groups which compared with Driscoll's in age and interval were two-year-olds retested three years later, three-year-olds retested two years later, and four-year-olds retested one year later.

In both studies subsequent Stanfords were higher than the Merrill-Palmers. The differences ranged from less than 1 point to 21 points in IQ. Correlations were similar in the two studies, with one exception being slightly higher from initial Kuhlmann to Stanford than from Merrill-Palmer to Stanford. The correlations from Merrill-Palmer ranged from .41 to .66 and from Kuhlmann .54 to .71.

Relation to Minnesota Preschool Scale

A small amount of work has been reported on comparisons of Merrill-Palmer results with the Minnesota Preschool Scale (Table 52). Jersild and Markey (14) give comparative means for two small groups of children. In one group the mean IQ on the Minnesota scale was 6 points higher than the Merrill-Palmer IQ; in the other group 3 points higher. In our group, which was small also, the means were similar on the two tests.

The correlation between Merrill-Palmer and Minnesota IQ's in our study was higher (.84 ± .04) than in Goodenough's (9) group, her correlation being .67. We obtained the same relationship (.77) be-

Table 51

Relation to Subsequent Stanford-Binets: Comparison of Merrill-Palmer and Kuhlmann-Binet

Investigator	Chronological Age, Years or Months	Scale	Children	Mean IQ	Stanford-Binet			Difference	r	P.E.
					Chronological Age, Years	Interval Months	Mean IQ			
Driscoll (7)	2 and 3	Merrill-Palmer	21	103.1	5 or over		121.7	18.6	.51 ± .09	
Driscoll (7)	2 and 3	Kuhlmann	39	119.8	5 or over		122.6	2.8	.66 ± .06	
Driscoll (7)	3 and 4	Merrill-Palmer	39	111.4	5 or over		123.3	11.9	.41 ± .09	
Driscoll (7)	3 and 4	Kuhlmann	54	123.0	5 or over		122.0	— 1.0	.71 ± .05	
Wellman	18 to 29	Merrill-Palmer	24	111.5		30 to 41	132.4	20.9	.66 ± .10	
Wellman	18 to 29	Kuhlmann	24	127.0		30 to 41	132.4	5.4	.62 ± .10	
Wellman	30 to 41	Merrill-Palmer	70	115.8		18 to 29	122.9	7.1	.49 ± .06	
Wellman	30 to 41	Kuhlmann	67	123.3		18 to 29	123.9	.6	.54 ± .06	
Wellman	42 to 53	Merrill-Palmer	75	122.2		6 to 17	129.5	7.3	.55 ± .05	
Wellman	42 to 53	Kuhlmann and Stanford	71	126.5		6 to 17	129.3	2.8	.59 ± .05	

Table 52

Relation to Minnesota Preschool Scale

Investigator	Children	Chronological Age, Months	Merrill-Palmer IQ	Minnesota IQ	Zero-Order Correlations				Mental Age (Chronological Age Constant)
					Mental Age	IQ	Verbal IQ	Nonverbal IQ	
Goodenough (9)	33	27 to 59			.90	.67			.72
Jersild and Markey (14)									
Group A	17	22 to 36	112.0	118.0					
Group B	19	33 to 50	125.0	128.0					
	18								
Wellman	27	29 to 60	106.7	104.9		.84 ± .04	.77 ± .05	.77 ± .06	

tween Merrill-Palmer and verbal IQ on the Minnesota scale as between Merrill-Palmer and the nonverbal Minnesota IQ.

Relation to Occupation of Father

Kawin (15) compared two groups of very different socio-economic status. One group of sixty-two children was composed of children attending the Winnetka nursery school. About 90 per cent of the fathers of these children were in the professional classes. The other group of sixty-two children was composed of children attending the Mary Crane nursery school, about 50 per cent of whose fathers were unskilled laborers. The two groups were paired on chronological age at the time of the Merrill-Palmer test, with a difference not exceeding two months between pairs.

Kawin found very little difference between these two groups in Merrill-Palmer scores. The Winnetka group made scores very slightly higher than the Mary Crane group, but the differences were not significant (Table 53). Their mental age was 2.2 months higher than that of the Mary Crane children, their sigma score .1 higher, their percentile rank 4.8 higher, and their IQ 5.7 higher. The ratios of the differences to the probable errors of the differences yielded the following: for mental age 1.5, for sigma score 1.3, and for percentile rank 1.4. The ratio for the difference between intelligence quotients was not given. These ratios indicate the following chances in 100 of true differences: 84, 81, and 83 respectively.

On Stanford-Binet, however, the Winnetka group was clearly superior to the Mary Crane group. Forty-eight of the Winnetka children and forty of the Mary Crane children received the Binet test. The Winnetka mean was 123.9 while the Mary Crane mean was 105.9, a difference of 18.0 points. This difference was statistically significant, the ratio of the difference to the probable error of the difference being 9.0.

Compared with Iowa City children entering the preschool laboratories, the Winnetka children scored low in Merrill-Palmer ability, but equal in Binet ability. About 75 per cent of the Iowa City group had fathers in the professional classes, as compared with about 90 per cent of the Winnetka children. The mean age was very similar, the Iowa City children being 1.9 months older. The Iowa City children were .53 sigma score superior to the Winnetka children, 19.0 percentile ranks superior, 8.3 points superior in Merrill-Palmer IQ, and 1.3

Table 53
Relation to Occupation of Father

Investigator	Group	Occupation	Children	Chronological Age, Months	Mental Age, Months	Sigma Score**	Percentile Rank	Intelligence Quotient	Binet Test: Children	Binet Test: Intelligence Quotient
Kawin (15)	Winnetka	About 90 per cent professional	62	36.7*	39.1	.2	57.0	106.5	48	123.9
Kawin (15)	Mary Crane	About 50 per cent unskilled laborers	62	36.6*	36.9	.1	52.2	100.8	40	105.9
Wellman	0 to 3 months attendance prior to test	About 75 per cent professional	111	38.6		.73	76.0	114.8	93	122.6
Wellman	0 to 3 months attendance prior to test	Professional	81	37.4		.68	73.6	113.7	65	122.7
Difference	Winnetka and Mary Crane			.1	2.2	.1	4.8	5.7		18.0
	Winnetka and Iowa City			— 1.9		— .53	—19.0	— 8.3		1.3
	Mary Crane and Iowa City			— 2.0		— .63	—23.8	— 14.0		— 16.7
Ratio of difference to probable error of difference	Winnetka and Mary Crane				1.5	1.3	1.4	***		9.0
Ratio of difference to standard deviation of difference	Winnetka and Iowa City					3.9	4.7			
	Mary Crane and Iowa City					4.3	5.5			

*Winnetka and Mary Crane children were paired on chronological age, with a difference not exceeding two months between pairs.

**Sigma scores in Kawin's study were interpolated to the nearest tenth whenever the score did not fall directly on the norms in Table 38 of Stutsman (27).

***Not given.

points inferior in Binet IQ. The ratios of the differences to the standard deviation of the differences were respectively: for sigma score 3.9; for percentile rank 4.7, indicating complete statistical significance. The ratio for Merrill-Palmer intelligence quotients could not be computed because Kawin did not report probable errors of these means. The difference in Binet IQ was obviously insignificant.

The portion of the Iowa City group whose fathers were in the professional classes made scores so similar to the total group of new entrants that significance of differences with Kawin's groups was computed only on the total group of new entrants.

Children entering the preschool laboratories were superior to the Mary Crane group as follows: .63 sigma score, 23.8 percentile ranks, 14.0 Merrill-Palmer IQ points, and 16.7 Binet IQ points. The ratio for sigma scores was 4.3 and for percentile ranks 5.5.

It is of interest that in the particular samplings involved neither Kawin nor the present study showed any substantial differences in Merrill-Palmer ability between different occupational classes and that both studies showed significant differences in Binet ability of the same children. Kawin's groups were much more widely separated occupationally than the groups compared in the present study. At the same time it is puzzling to find an explanation for the low Merrill-Palmer performance of the Winnetka children in comparison with Iowa City children. It does not appear to be in such factors as conditions of testing and training of examiners. Iowa City children live in a small community, with ample outdoor and indoor play space. In general they are provided with a well-selected assortment of play materials at home, and the direct responsibility for care of the children falls upon the mother who is vitally interested in her child's development. These may be factors affecting the intellectual level of the children when they enter preschool.

Relation to Other Characteristics of Children

Correlations between Merrill-Palmer scores and a number of characteristics were computed by Goodenough (9) on a group of thirty-three nursery school children. These correlations have been rearranged in Table 54 in order of size of relationship with Merrill-Palmer IQ. The highest correlations for IQ (.46, .42, and .41) were with length of nursery school attendance, ratings on attractiveness of personality, and ratings on beauty. Other positive relations were with leadership,

Table 54
Relation to Other Characteristics: Comparison with Minnesota Preschool Scale Results on Thirty-Three Nursery School Children, Twenty-Seven to Fifty-Nine Months of Age*

Characteristics	Mental Age				Intelligence Quotients	
	Zero Order Correlations		Chronological Age, Constant		Zero Order Correlations	
	Merrill-Palmer	Minnesota	Merrill-Palmer	Minnesota	Merrill-Palmer	Minnesota
Length of nursery school attendance	.74	.73	.42	.44	.46	.30
Attractiveness of personality	.15	.16	.58	.49	.42	.23
Beauty	.25	.19	.52	.34	.41	.29
Leadership	.66 .72	.64 .74	.15	.20	.36	.10
Sociality	.62	.83			.35	.08
Talkativeness	.46	.49	.33	.37	.35	.32
Marston scale of introversion-extroversion	.16	.14	.21	.15	.20	.03
Laughter	.40	.41	.23	.11	.13	.07
Physical activity	.29	.21	.09	—.03	.11	—.07
Compliance	—.35	—.34	—.17	—.17	—.27	—.15
Olson scale for prediction of behavior problems	—.34	—.41	—.50	—.55	—.29	—.32
Height	.73	.74	.03	.23		
Weight	.46	.44	—.10	—.04		

*Adapted from Goodenough (9).

sociality, and talkativeness (.36, .35, and .35). Two negative correlations are of interest: compliance and Olson's scale for prediction of behavior problems (—.27 and —.29). Children with high Merrill-Palmer IQ's thus tended not to be conformers but were not behavior problems.

The correlations with Minnesota Preschool Scale IQ tended to line up in rank approximately the same as those with Merrill-Palmer IQ, although the relationships were lower. Correlations with mental age when chronological age was held constant agreed with the IQ correlations in finding the highest relationship with length of nursery school attendance, attractiveness of personality, and beauty and in finding negative correlations with compliance and behavior problems.

Markey (19) computed correlations with imaginative behavior as measured in three situations: free play, block building, and a housekeeping game. These are presented in Table 55. The score on imaginative behavior in the free play situation correlated .39 ± .09 with Merrill-Palmer IQ and .64 ± .06 with mental age. The correlation based on median mental age on Merrill-Palmer, Kuhlmann-Binet and

Table 55

Relation to Imaginative Behavior*

Traits Correlated	Children	Chronological Age, Months	Merrill-Palmer Mental Age	Merrill-Palmer IQ	Median Mental Age on Merrill-Palmer, Minnesota and Kuhlmann	Kuhlmann			Minnesota		
						Children	Mental Age	IQ	Children	Mental Age	IQ
Free Play											
Total imaginative behavior	46	26 to 50	.64 ± .06	.39 ± .09	.57 ± .07						
Imaginative situations in which child participated	46	26 to 50	.68 ± .05	.41 ± .08	.62 ± .06						
Original items of overt behavior and language	46	26 to 50	.61 ± .06	.36 ± .09	.56 ± .07						
Imaginative situations initiated by child	46	26 to 50	.60 ± .06	.40 ± .08	.51 ± .07						
Group imaginative situations which child initiated	46	26 to 50	.62 ± .06	.40 ± .08	.56 ± .07						
Block Building											
Items constructed	75	18 to 59	.58	.24 ± .07		62	.64	.29 ± .08	96	.65	.28 ± .06
Names assigned to constructions	75	18 to 59	.51	.07 ± .08		62	.69	.42 ± .07	96	.59	.21 ± .07
Items constructed and names assigned	75	18 to 59	.60	.21 ± .07		62	.71	.40 ± .07	96	.71	.35 ± .06
Items of overt imaginative behavior and language	75	18 to 59	.61	.23 ± .07		62	.67	.35 ± .08	96	.69	.31 ± .06
Housekeeping Game											
Items of overt imaginative behavior and language	24	24 to 59	.29	.16		21	.49	.38	41	.27	.15

*Adapted from Markey (19).

Minnesota Preschool Scale was .57 ± .07. Subscores of imaginative behavior in free play yielded correlations similar to the total score.

Imaginative behavior in the block-building situation showed little relationship to Merrill-Palmer IQ the correlations ranging from .07 to .24 ± .07. With Minnesota IQ the relationships were slightly higher, ranging from .21 ± .07 to .35 ± .06, and with Kuhlmann-Binet IQ again slightly higher, ranging from .29 ± .08 to .42 ± .07.

The results of this study indicate that there is a slight positive relationship between Merrill-Palmer ability and imaginative behavior, and that the degree of relationship is somewhat dependent on the situation used in measuring imaginative behavior.

Cushing (6) reported correlations of perseverative tendency in preschool children as measured by six tests with Merrill-Palmer mental age, Binet mental age, and Binet IQ. The correlations for forty-nine children 23 to 60 months of age were: with Merrill-Palmer mental age .58, with Binet mental age .48, and with Binet IQ .25.

Nelson (22) in studying resistant behavior in a mental test situation found that children who were below the median in resistance to Merrill-Palmer test items were higher in mental age and intelligence quotient than children who were above the median in resistance. The respective mental ages were 52.2 and 41.9 months and the mean IQ's were 125.3 and 109.0.

Prediction From Infant Tests

Hubbard (13) presented correlations for two small groups on prediction of Merrill-Palmer ability from the Bühler infant test. The first group was composed of twenty-five children tested on the Merrill-Palmer at 16 to 37 months of age, mean age 23.4 months, who had received the Bühler test 1 to 26 months earlier, mean interval 13.2 months. The correlation between "ratings" (scoring used not specified) on the two tests was .37 ± .11. When corrected for attenuation, the correlation was .52. Fifteen children received a second Bühler test. The correlation with Merrill-Palmer 10 months later was .70 ± .01. Corrected for attenuation, the correlation became .99. The above correlations were reported as between "ratings" on the two scales, without statement as to whether the rating was raw score, mental age, sigma score, IQ, or percentile rank.

Comparison With English Norms

Bristol (4) applied the Merrill-Palmer test to 531 English children ages twenty-four to seventy-one months. She worked out new mental age equivalents for raw scores. In the following tabulation the mental age equivalent for every tenth score is reproduced and compared with Stutsman's mental ages:

Raw Score	Mental Age Equivalent, Months: Stutsman (27)	Bristol (4)	Difference
0	18 or less	18 or less	0
10	22	20	— 2
20	28	25	— 3
30	32	30	— 2
40	36	34	— 2
50	40	38	— 2
60	47	43	— 4
70	54	48	— 6
80	63	57	— 6
90	78	67	—11
93	78 or more	70	— 8

It will be seen that Bristol's mental ages were consistently lower than Stutsman's and that the difference becomes greater with increase in score. Bristol's children were drawn from the following sources: private homes 112 children, nursery schools 253, day nurseries 5, infant schools 72, convalescent home 11, child guidance clinic 42, and kindergartens 36. She states that they represented all classes and all occupations and were equally distributed as to sex.

Bristol also presented a table of percentile ranks in terms of score values. In the following tabulation the scores that received percentile ranks of 10, 50, and 90 at exact half-year ages in months are reproduced and compared with Stutsman's figures:

Chronological Age, Months	Stutsman	Bristol	Difference
	10th Percentile		
24	7	10	3
30	13	18	5
36	24	28	4
42	36	39	3
48	44	49	5
54	55	59	4
60	63	66	3
66		73	
	50th Percentile		
24	13	17	4
30	25	30	5
36	40	43	3
42	52	56	4
48	62	69	7
54	70	77	7
60	77	83	6
66		87	

	90th Percentile		
24	21	25	4
30	36	42	6
36	54	58	4
42	67	70	3
48	75	80	5
54	82	86	4
60	89	90	1
66		93	

It will be seen that a higher raw score was required for a given percentile rank according to Bristol than Stutsman provides. For example a percentile rank of 90 was assigned a child of 24 months of age who made a raw score of 21 in Stutsman's standardization, but a score of 25 was required by Bristol.

Summary

The results of other investigators have been brought together in this section and our results compared with theirs.

Relation to Chronological Age

Correlations between mental age and chronological age ranged from .81 ± .01 to .91 ± .01 (Bristol, Goodenough, Wellman), indicating a very high relationship. There was no correlation between chronological age and IQ (Goodenough, Wellman). Raw scores correlated with chronological age .66 to .92 ± .004 (Kawin, Stutsman).

Sex Differences

Girls usually made slightly higher percentile scores than boys (Stutsman, Wellman).

Gain According to Interval

1. Significant gains in IQ over six- and twelve-month intervals were found by Driscoll and Wellman. Correlations were strikingly similar, in both studies decreasing with the longer interval (.62 and .63 at six months, .47 and .52 at twelve months).

2. There were gains in sigma scores over a six-month interval between first and second tests, but no gain in a comparable interval between second and third tests (Kawin, Wellman). In our group the differences appeared to be a function of the amount of intervening preschool attendance.

Correlations between sigma scores on retests ranged from .47 ± .01 to .92 ± .02 (Kawin, Stutsman, Wellman). Most of the correlations fell between .47 and .59. There was a decrease in the correlation with increase in interval.

3. Binet correlations were higher than for sigma scores (Kawin,

Wellman). Kawin's correlations (.75 ± .05 to .84 ± .03) were higher than ours (.61 ± .07 to .69 ± .06).

4. Correlations between raw scores were, as would be expected, higher than between IQ's or sigma scores. The range was from .67 ± .04 to .91 ± .01 (Kawin, Stutsman). Narrowing the age range reduced the correlations.

Age Differences on Retests

1. Gains in IQ were made by each age group, with one exception (Driscoll, Wellman). There was no trend for the correlation between IQ's to increase or decrease with age (Driscoll). For the total age range, the correlations over a six-month interval were practically identical in the two studies (.62 ± .05 and .63 ± .04).

2. There was an increase in the mean IQ with age (Driscoll, Wellman). In both studies the mean IQ was at the average level at two years and very superior at four years of age.

3. Binet correlations ran higher than Merrill-Palmer correlations, age by age, at six-month intervals and rose with increasing age (Driscoll). However, when the total age range was included, the correlations for Binet and Merrill-Palmer were the same (Driscoll, Wellman). Driscoll's correlation compared more closely with ours covering a period of preschool attendance than with nonattendance. When the interval between tests was twelve months, Binet correlations were higher than Merrill-Palmer (Driscoll, Wellman). Driscoll's and our results are in close agreement (.47 and .52 for Merrill-Palmer, .60 and .63 for Binet).

Sex Differences in Gains

Sex differences in the percentage gaining, median gains, and median losses were inconsistent in the studies by Driscoll and Wellman. The two studies agreed that a smaller percentage of children gained on Binet than on Merrill-Palmer. When Merrill-Palmer comparisons showed differences between the sexes in percentage gaining, there was the same direction of sex difference on the Binet. It appeared that whatever induced gains on the one scale also affected the other.

Relation of Gains to IQ Level

There was an inverse relationship between percentage gaining and IQ level (Driscoll, Wellman). Binet changes showed the same tendency. Median gains and median losses on the Merrill-Palmer were irregular in amount at the different IQ levels, but Binet changes were

regular, showing an increase in median gain and a decrease in median loss inversely to IQ level.

Effect of Preschool Attendance

1. All studies that have compared gains during periods of preschool attendance and nonattendance agree in gains made by children attending preschool (Barrett and Koch, Kawin and Hoefer, Waring, Wellman). The amount of gain at a given IQ level is in close agreement in the different studies. All studies show gains also during nonattendance, but the gains were less than during attendance, except in the study by Kawin and Hoefer, where the control group gained as much as the preschool group. The other three studies are in close agreement on the amount gained during nonattendance.

2. Goodenough's correlation between IQ and number of days of preschool attendance was higher than the correlations obtained in our study. Our study gave conflicting results according to sex and according to method of scoring used.

Relation to Binet

1. Higher mental ages were found on the Binet than on the Merrill-Palmer (Cushing, Rust, Wellman).

2. Correlations between Merrill-Palmer and Binet mental ages ranged from .63 ± .04 to .94 ± .01 (Kawin, Kuenzel, Rust, Stutsman).

3. Higher IQ's were found on the Binet than on the Merrill-Palmer (Driscoll, Jersild and Markey, Nelson, Rust, Wellman).

4. Correlations between Merrill-Palmer and Binet IQ's ranged from .45 ± .05 to .76 ± .05 (Driscoll, Nelson, Rust, Wellman).

5. While the children in our study showed a definite trend for the difference between Merrill-Palmer and Binet IQ to vary with IQ level, Rust's group did not show this trend. Driscoll's results are, however, in close agreement with ours in showing an increasing divergence between IQ's on the two tests as expressed in percentage of children for whom the Merrill-Palmer was higher or lower and median amounts of difference.

6. In both Driscoll's and our study subsequent Stanford-Binets were higher than the Merrill-Palmers. Correlations were similar in the two studies, tending to run slightly lower between Merrill-Palmer and Stanfords (.41 to .66) than between initial Binets and later Binets (.54 to .71).

Relation to Minnesota Preschool Scale

Only a small amount of work, with small numbers of cases, has been done on comparisons with the Minnesota preschool scale.

Relation to Occupation of Father

Neither Kawin's study nor our study showed any substantial difference in Merrill-Palmer ability between different occupational classes, although both studies showed significant differences in Binet ability in favor of the higher occupations. Kawin's children from the professional occupations (Winnetka group) were significantly lower in Merrill-Palmer ability than our children, but were equal to them in Binet ability.

Relation to Other Characteristics

Positive correlations between Merrill-Palmer IQ and length of nursery school attendance, attractiveness of personality, beauty, leadership, sociality, and talkativeness (.35 to .46) were reported by Goodenough. Slight negative correlations were found with compliance and with the Olson scale for the prediction of behavior problems (—.27 and —.29).

Positive correlations with imaginative behavior were reported by Markey, with the degree of relation dependent upon the measure used. A positive correlation of mental age with perseverative tendency was reported by Cushing. Nelson found that children who showed more resistance to the Merrill-Palmer test were lower in mental age than less resistant children.

Prediction From Infant Tests

A positive correlation between the Bühler infant tests and Merrill-Palmers about one year later for a small group of children has been reported by Hubbard.

English Norms

Bristol found that British children made higher scores than the children in Stutsman's standardization group. New norms for English children were worked out.

SUMMARY OF RESULTS AND CONCLUSIONS AND AN EVALUATION OF THE TEST

The data upon which this study is based consist of 510 Merrill-Palmer tests given to 281 children. The results on the Merrill-Palmer scale are compared throughout to Kuhlmann- and Stanford-Binet tests of the same children. The relation of Merrill-Palmer ability to subsequent Binets is also considered. Comparisons are made with the results of other investigators.

A summary of results is presented at the end of each section of this study. The reader is referred to these summaries for details of the findings. A few of the conclusions of more general interest have been selected for reiteration here.

Persons using the Merrill-Palmer scale must decide upon which of the scoring methods (mental age, sigma score, intelligence quotient, percentile rank) will be used in analyses. The results of the present study indicate that intelligence quotient is as useful and adequate a score as sigma score (standard deviation in terms of raw score) or percentile rank. There appear to be some inadequacies in the original standardization group and in the method of deriving scores from the raw scores. So far as the present study contributes anything to this problem, it would appear that the difficulty is not restricted to any one derived score but applies to all.

One fact that emerges rather clearly from the present study and from the results of other investigators is that gains may generally be expected on retests. These gains have been variously explained by different investigators, most frequently as practice effects or unreliability of the test. The present investigation has contributed to this problem by demonstrating that the pattern of change on retest at one week was quite different from the pattern of change over a longer period, such as six months. Amount of change in a one-week interval was positively correlated with initial IQ, indicating that the brighter children made the greatest gains, while at a six-month interval amount of change was negatively related to initial IQ, indicating the greatest gains were being made by the duller children. Over the short interval all levels gained, while over the longer period the brightest children lost in IQ. From these different patterns of change, it was inferred that the causes of change were not the same over the longer and shorter periods. The best explanation of gains at one week appeared to be practice (memory) effects. It seems very reasonable to expect better

performance on retest since so many of the tests are of the form board variety and scored in terms of speed. However, straight practice effects do not constitute an adequate explanation of the gains made over the longer period. Neither do such explanations as statistical regression or biological regression fit both cases. In some instances gains occurred over the longer interval which it seemed illogical to credit as real increases in ability because of the relatively barren environment of the children. Whatever the causes, the results show that an increase of approximately five points in IQ may be expected on retest at an interval of six months.

Children attending the preschool laboratories at Iowa City were found to make gains in IQ over and above the gains to be expected from repetition of the test. Gains made over the winter months when attending preschool were approximately twice the size of the gains made over the summer vacation months, the interval being the same. It was felt that the difference in gains could be attributed to the effects of preschool attendance. Other types of analysis bore out the general conclusion that preschool attendance materially affected the ability of these children, although subdivisions within a preschool year, such as in number of days attendance, appeared to be too fine to reveal differential effects. The changes in Merrill-Palmer ability were less cleancut than in Binet ability.

Neither our results nor the results of other investigators have revealed significant differences between children from different occupational classes. Our group was a highly selected one, since no children attended preschool who were below occupational Class IV. However, Binet IQ's of the same children showed significant differences in favor of the children from the professional classes.

No outstanding differences were revealed in Merrill-Palmer ability between children whose parents were better educated and those whose parents were less well educated. Our group was highly selected, comparisons being made between two parent groupings, eight to fifteen years of schooling, and sixteen to twenty years of schooling. The lack of difference in Merrill-Palmer ability is interesting in contrast with a difference in Binet ability which closely approached statistical significance when midparent education was compared.

Sufficient correspondence between Merrill-Palmer and Binet ability was found to justify the conclusion that the two tests measure something in common. However, the correlations were low as between two

tests of intelligence, ranging from about .45 to .65 for IQ on simultaneous tests (interval one week). Analysis of the characteristics of divergence between the two tests suggests that the two tests measure somewhat different types of abilities.

Of considerable interest is the problem of prediction of later ability from a test at the preschool ages. The present study contributes to this problem only in so far as later Binet ability is concerned. Correlations of Merrill-Palmer with subsequent Binets at intervals varying from six months to six years indicated that the relationship was slightly lower than from early Binet to subsequent Binet. The median correlation for Merrill-Palmer was .44 as compared with .53 for early Binets. The range for Merrill-Palmer was .34 to .66 and for early Binet .33 to .79. When an average IQ of Merrill-Palmer and early Binet was correlated with subsequent Binets, the median correlation was .57. The range was .41 to .81. Adding the Merrill-Palmer to early Binet thus did not materially affect the relationship of early Binet to subsequent Binet. It should be kept in mind in evaluating these correlations that correlations covering periods of preschool attendance were lower than those covering the same time period of nonattendance. In other words, preschool attendance served to reduce consistency between test and retest. Since a large proportion of the group given subsequent Binets attended preschool, the actual size of the correlations should be interpreted accordingly. Low correlations on retest should not be interpreted as entirely due to unreliability of the test.

It is of interest that when combined Merrill-Palmer and Binet IQ's were correlated with combined later IQ's, the correlations were higher than for either test alone on combined age groups, being .71 $\pm$ .05 for a six-month interval and .72 $\pm$ .06 for a twelve month interval. These, however, were within the range of correlations for separate age groups from early Binet alone. This procedure, of course, can be used only within the age limits to which the Merrill-Palmer is applicable.

In our study girls tended to receive higher scores than boys.

Comparison with the results of other investigators showed surprisingly close agreement in most respects. Two points of major disagreement may be mentioned here. Kawin and Hoefer in a study of the effects of preschool attendance found equal gains for preschool and control groups, while our results showed approximately double the gain in IQ when the children were attending preschool than when not attending. By comparison with the results of other investigators,

it was shown that Kawin and Hoefer's control group far exceeded any other control group in amount of gain. The results of Barrett and Koch and of Waring are in close agreement with ours in tendency and in amount of change.

Another point of disagreement is the divergence between Merrill-Palmer and Binet per IQ level. Rust's group of three-year-old children showed no particular trend for the amount of difference between the two tests to vary at the different brightness levels. In our group the divergence increased with successively higher IQ classifications and successively lower classifications from a given point whether classified on Binet or on Merrill-Palmer. Driscoll's group was in agreement with ours. The reason for the disagreement with Rust is not apparent. Rust's results indicate that the divergence found in our group should not hastily be ascribed to a statistical necessity due to regression. Possibly the environment of Rust's children gave relatively more emphasis to the types of abilities measured by the Binet test, or stated conversely, possibly the environments of our children emphasized more the types of abilities measured by the Merrill-Palmer scale. The majority of Rust's group was not attending nursery school.

Summing up the various points, it appears that the Merrill-Palmer test is not quite as adequate a test for the preschool ages as the Binet test. However, it appears to measure a different set of abilities to a considerable extent, and therefore might be retained as a supplementary test. The relatively low agreement with a simultaneous Binet indicates that it cannot be substituted for the Binet with any great assurance of comparable results. Whether the Merrill-Palmer presents a continuum with performance scales at later ages remains to be demonstrated, there being no evidence available at this time. A point decidedly in favor of the Merrill-Palmer test is that children and examiners like it. It appeals to children and usually requires less skill on the part of the examiner to secure full co-operation of the child than is required by the more verbal type of test. However, ease of administration does not necessarily make a test more adequate. It is the writer's opinion that one of the difficulties with the Merrill-Palmer scale is its reliance on small differences in speed of performance for crediting the performance at different levels. In some instances the absurdly small difference of one or two seconds determines what credit the test will receive. Objection to discrimination on such small margin is twofold, on the reaction time of examiner and on motivation of the young child to speed. While it seems unfortunate that speed re-

ceives so much emphasis in scoring, the general adequacy of the test despite such apparent weaknesses must not be overlooked. Whether the test would be improved by minimizing the speed factor can be determined only by experimentation. Greater attention to comparable selection of the standardization group at the different ages might overcome some of the present discrepancies. In its present form, the adequacy of the scale will need to be judged by each person using it in terms of the uses to which the results are to be put. It is hoped that the present study will be an aid in such evaluation.

REFERENCES

1. Aldrich, Cecelia G., and Doll, Edgar A.: Comparative intelligence of idiots and normal infants. Ped. Sem. & J. Genet. Psychol., 1931, 39, 227-257.
2. Barrett, Helen Elizabeth, and Koch, Helen Lois: The effect of nursery school training upon the mental test performance of a group of orphanage children. Ped. Sem. & J. Genet. Psychol., 1930, 37, 102-122.
3. Bridges, K. M. Banham: Critical notes on mental tests for children of preschool age. Ped. Sem. & J. Genet. Psychol., 1927, 34, 38-44.
4. Bristol, Hilda: An English norm for the Merrill-Palmer performance test: Based on a study of 530 children between the ages of two and six years. Brit. J. Educ. Psychol., 1936, 6, 250-266.
5. Coffey, Hubert S., and Wellman, Beth L.; The role of cultural status in intelligence changes of preschool children. J. Exper. Educ., 1936-1937, 5, 191-202.
6. Cushing, Hazel Morton: A perseverative tendency in preschool children: A study in personality differences. Arch. Psychol., 1929-1930, 17, No. 108, Pp. 55.
7. Driscoll, Gertrude Porter: The developmental status of the pre-school child as a prognosis of future development. Teach. Coll., Columbia Univ., Child Develop. Monog., 1933, No. 13, Pp. xi, 150.
8. Goodenough, Florence L.: The Kuhlmann-Binet tests for children of preschool age: A critical study and evaluation. University of Minnesota, Institute of Child Welfare Monograph Series, No. 2. Minneapolis, Minn.: University of Minnesota Press, [c. 1928] Pp. 146.
9. Goodenough, Florence L,: Inter-relationships in the behavior of young children. Child Develop., 1930, 1, 29-47.
10. Gordon, R. G.: The Merrill-Palmer scale of intelligence tests for preschool children applied to low-grade mental defectives. Brit. J. Psychol., 1933-1934, 24, 178-186.
11. Hatt, Elise: A comparative study of measurements and qualitative ratings of nursery school children. Child Develop., 1931, 2, 1-29.
12. Herring, Amanda, and Koch, Helen Lois: A study of some factors influencing the interest span of preschool children. Ped. Sem. & J. Genet. Psychol., 1930, 38, 249-279.

13. Hubbard, Ruth M.: A study of the reliability and validity of the Bühler infant scale. Ped. Sem. & J. Genet. Psychol., 1935, 47, 361-384.
14. Jersild, Arthur T., and Markey, Frances V.: Conflicts between preschool children. Teach. Coll., Columbia Univ., Child Develop. Monog., 1935, No. 21, Pp. ix, 181.
15. Kawin, Ethel: Children of preschool age: Studies in socio-economic status, social adjustment and mental ability, with illustrative cases. Chicago, Ill.: University of Chicago Press, 1934. Pp. xxv, 340.
16. Kawin, Ethel, and Hoefer, Carolyn: A comparative study of a nursery-school versus a non-nursery-school group. Chicago, Ill.: University of Chicago Press, 1931. Pp. ix, 52.
17. Kuenzel, M. W.: Mooseheart mental growth study. (The Mooseheart laboratory for Child Research) Abstract 736. Child Develop. Abstracts, 1934, 8, 236-237.
18. Kuhlmann, Frederick: A handbook of mental tests: A further revision and extension of the Binet-Simon scale. Baltimore, Md.: Warwick & York, 1922, Pp. 208.
19. Markey, Frances V.: Imaginative behavior of preschool children. Teach. Coll., Columbia Univ., Child Develop. Monog., 1935, No. 18, Pp. xvi, 139.
20. Mohr, George J., and Bartelme, Phyllis: Mental and physical development of children prematurely born. Amer. J. Dis. Child., 1930, 40, 1000-1015.
21. Mowrer, Willie Mae C.: Performance of children in Stutsman tests. Child Develop., 1934, 5, 93-96.
22. Nelson, Janet Fowler: A preliminary report on some uses of the psychological test situation for studying personality difference. [In] Thomas, Dorothy S., and associates: Some New Techniques for Studying Social Behavior. Teach. Coll., Columbia Univ., Child Develop. Monog., 1929, No. 1, Pp. x, 203. (p. 118-149).
23. Nelson, Janet Fowler: Personality and intelligence: A study of some responses other than intellectual noted in a simple mental test situation. Teach. Coll., Columbia Univ., Child Develop. Monog., 1931, No. 4, Pp. 62.
24. Nemzek, Claude L.: The constancy of the I.Q. Psychol. Bull., 1933, 30, 143-168.
25. Rust, Metta Maund: The effect of resistance on intelligence test scores. Teach. Coll., Columbia Univ., Child Develop. Monog., 1931, No. 6, Pp. xi, 80.
26. Stutsman, Rachel: Performance tests for children of pre-school age. Genet. Psychol. Monog., 1926, 1, 1-67.
27. Stutsman, Rachel: Mental measurement of preschool children: With a guide for the administration of the Merrill-Palmer scale of mental tests. Yonkers-on-Hudson, N. Y.: World Book Co., [c. 1931] Pp. x, 368.
28. Stutsman, Rachel: Factors to be considered in measuring the reliability of a mental test, with special reference to the Merrill-Palmer scale. J. Educ. Psychol., 1934, 25, 630-633.

29. Terman, Lewis M.: The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon intelligence scale. Boston, Mass.: Houghton Mifflin, [c. 1916] Pp. xviii, 362.
30. Terman, Lewis M., and Merrill, Maude A.: Measuring intelligence: A guide to the administration of the new revised Stanford-Binet tests of intelligence. Boston: Houghton Mifflin, [c. 1937] Pp. xiv, 461.
31. Thurstone, L. L.: The vectors of mind. Psychol. Rev., 1934, 41, 1-32.
32. Thurstone, L. L.: A new conception of intelligence. Educ. Rec., 1936, 17, 441-450.
33. Vance, Thomas F.: The food selections of nursery school children. Proc. Iowa Acad. Sci., 1930, 37, 343-350.
34. Vance, Thomas F., Prall, O. L., Simpson, Ruby V., and McLaughlin, Florence E.: The development of children in the home management houses of the Iowa State College. J. Exper. Educ., 1933-1934, 2, 166-169.
35. Wagoner, Lovisa C., and Armstrong, Edna M.: The motor control of children as involved in the dressing process. Ped. Sem. & J. Genet. Psychol., 1928, 35, 84-97.
36. Waring, Ethel B.: A report of the psychological examinations of preschool children in the nursery school over a period of eight years. [In] Ten-year Report of Studies in Child Development and Parent Education, Cornell Univ. Agric. Experiment Station, Cont. from Stud. in Home Econ. Bull., 1935, No. 638, 42-43.
37. Wellman, Beth L.: The effect of pre-school attendance upon the IQ. J. Exper. Educ., 1932-1933, 1, 48-69.
38. Wellman, Beth L.: Growth in intelligence under differing school environments. J. Exper. Educ., 1934-1935, 3, 59-83.
39. Wellman, Beth L.: The permanence of early training effects: A longitudinal study of 41 subjects from preschool through senior high school and 41 subjects from the primary grades through senior high school. [In] Proceedings, Second Biennial Meeting of the Society for Research in Child Development, 1936. Pp. xv, 166.
40. Wellman, Beth L.: Mental growth from preschool to college. J. Exper. Educ., 1937-1938, 6, 127-138.
41. Wilson, Charles A., Sweeny, Mary E., Stutsman, Rachel, Chesire, Leona E., and Hatt, Elise.: The Merrill-Palmer standards of physical and mental growth. Detroit, Mich.: Merrill-Palmer School, 1930. Pp. ix, 121.

S University of Iowa-Studies in Child Welfare-Vol.XV-No.3
(Wellman, Beth L.)
The Intelligence of Preschool Children as Measured by the

S UNIVERSITY OF IOWA

Studies in Child Welfare
Vol.XV, No.3

Wellman, Beth:
The Intelligence of Preschool Children as Measured by the Merrill-Palmer Scale of Performance Tests.

ONTARIO COLLEGE OF EDUCATION
DEPT. OF EDUCATIONAL RESEARCH
UNIVERSITY OF TORONTO

Obtainable from the Department of Publications

Paper bound ----------$1.00
Cloth bound ---------- 1.35

R.W.B. JACKSON LIBRARY
3 0005 03044701 8

155.4
I64U
v.15, no.3
Iowa. University
University of Iowa studies in child welfare

155.4
I64U
v.15, no.3
Iowa. University
University of Iowa studies in child welfare